
Fluorescence spectroscopic studies of protein
conformational dynamics

Dissertation
for the award of the degree
"Doctor rerum naturalium"
of the Georg-August-Universität Göttingen

within the doctoral program GGNB
of the Georg-August University School of Science (GAUSS)

submitted by
Phillip Gunther Kroehn
from Hannover

Göttingen 13.08.2013

Members of the Thesis Committee:

Prof. Dr. J. Enderlein (Reviewer)

3rd Institute of Physics - Biophysics, Faculty of Physics,
Georg-August-University Göttingen

Prof. Dr. H. Grubmüller (Reviewer)

Department of Theoretical and Computational Biophysics
Max Planck Institute for Biophysical Chemistry

Prof. Dr. A. Janshoff

Institute of Physical Chemistry, Faculty of Chemistry
Georg-August-University Göttingen

Date of oral examination: 21.10.2013

Affidavit

Hereby, I declare that the presented thesis has been written independently and with no other sources and aids than quoted.

Göttingen, 13.08.2013

Contents

Abstract	1
1 Introduction	3
1.1 The cell	3
1.2 Proteins	3
1.3 The structure of proteins	4
1.4 Protein folding	6
1.5 Beta sheets and WW-domains	8
1.6 Goal of this project	11
2 Materials and Methods	13
2.1 Materials	13
2.2 Plasmid DNA mini preparation	13
2.2.1 Purification of DNA	13
2.2.2 Determination of nucleic acid concentration	14
2.2.3 Restriction of nucleic acids	14
2.2.4 Agarose gel electrophoresis	14
2.2.5 DNA extraction from agarose gels	15
2.2.6 Ligation of DNA	16
2.3 Transformation of nucleic acids	16
2.4 Polymerase chain reaction (PCR)	16
2.4.1 QC Polymerase chain reaction	17
2.4.2 Primer design for QC PCR	17
2.4.3 DpnI digestion	18
2.5 <i>E. coli</i> cell culture	18
2.5.1 Bacterial strains and DNA plasmid vectors	18
2.5.2 LB media	19
2.5.3 Preparation of competent cells	19
2.5.4 Protein Expression	20
2.6 Purification of recombinant proteins	20
2.6.1 Protein purification by nickel-NTA	20
2.6.2 Protein purification by glutathione sepharose	21
2.6.3 Thrombin proteolysis	21
2.6.4 Size exclusion chromatography by HPLC	22

2.7	Determination of protein concentration	22
2.8	Electrophoresis	23
2.8.1	SDS-PAGE	23
2.8.2	Staining of SDS-PAGE	25
2.9	Mass spectrometry of proteins	26
2.10	Nano-containers	26
2.11	Circular dichroism (CD)	29
2.11.1	Chiral molecules	29
2.11.2	Theory of circular dichroism	29
2.11.3	Circular dichroism measurement	31
2.12	Differential scanning calorimetry (DSC)	32
2.13	FRET	34
2.13.1	Fluorescence	34
2.13.2	Theory of FRET	37
2.13.3	Selection of FRET pairs	41
2.13.4	Acceptor labeling	43
2.13.5	Ion-exchange purification	44
2.13.6	Donor labeling	44
2.13.7	Measurements in the fluorescence photometer	45
2.13.8	Single molecule spectroscopy	45
2.13.9	Setup	49
2.13.10	Single molecule FRET experiments	50
3	Results: Preparation of WW-domain FRET sample	53
3.1	WW-domain purification by His-tag	53
3.2	Calmodulin-WW-domain	60
3.3	GST-hPin1-WW-domain	63
3.3.1	SDS-PAGE	66
3.3.2	HPLC purification	67
3.3.3	Determination of protein concentration	69
3.3.4	Mass spectrometry of proteins	70
3.3.5	Acceptor labeling	71
3.3.6	Ion-exchange purification	72
3.3.7	Donor labeling	74
3.4	Discussion	76

4	Results: Measurement of WW-domain FRET sample	77
4.1	Circular dichroism (CD)	77
4.2	Differential scanning calorimetry (DSC) of the hPin1-WW-domain.	78
4.3	FRET experiments of hPin1-WW-domain in bulky solutions	79
4.4	Single molecule FRET experiments of the hPin1-WW-domain	81
4.5	Nano-containers	83
4.6	Discussion	85
5	Conclusion and outlook	87
	Appendix	89
	<i>Vectors</i>	89
	Bibliography	91
	List of Figures	97
	Acronyms	100
	Acknowledgments	102
	Curriculum vitae	104

Abstract

The folding of a protein into its native structure is the fundamental prerequisite for its functionality. The polypeptide chain itself contains the chemical information necessary for spontaneous folding of the protein. Nevertheless, many proteins need chaperones in order to support their folding into their native structure. It is currently possible to predict the three dimensional structure of some small proteins from their primary structure. Linking protein structure and chain composition, however, is essential for understanding of molecular mechanisms in every living organism. Detailed understanding of the protein folding process is one of the most important questions in the natural sciences.

Protein folding occurs in microseconds to minutes on nanometer length scales. Therefore, a method of high accuracy is necessary in order to monitor protein folding.

Förster resonance energy transfer (FRET) is a method that offers the unique opportunity to measure distances in molecules with high precision. Application of single molecule experiments is the only way that allows monitoring folding events in detail. In this work, protein folding is measured using single molecule FRET experiments.

The diffusion of small molecules is very fast (around 10^{-6} cm²/s) and therefore challenging to determine. In the present study, this problem is addressed by using a highly precise setup specialized for single molecule measurements that allows for determining diffusion coefficients with an accuracy of 10% or better.

A WW-domain protein is used as a model system for the folding of small globular proteins and therefore yields fundamental insights into the principle of protein folding in general. WW-domains are small, three-stranded, all- β protein-domains (6 kDa) which appear in several different proteins.

In general, the handling of small proteins with standard biochemical methods is very challenging. However, in this work, we show how the preparation of WW-domain FRET samples was conducted combining several protein purification techniques and, thus, overcoming the difficulties of labeling.

1 Introduction

1.1 The cell

All organisms consist of cells, with the most simple species consisting of a single one (shown in figure 1). A cell is a body that is enclosed by a membrane which constitutes a barrier to close the cell off from the outside world.

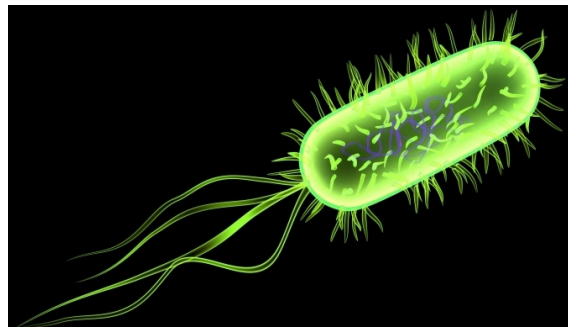


Figure 1: Schematic drawing of an E. coli cell (taken from: ishbytes (^[1]))

This barrier consists mainly of phospholipids. The barrier creates controlled conditions that are not in equilibrium with the external environment of the cell. A fundamental property of cells is their ability to reproduce themselves, i.e. to autonomously generate an identical copy. All the information necessary for this is contained in the genome of the cell. There, **deoxyribonucleic acid** (DNA) molecules serve as the cell's data storage. The information coded in DNA molecules helps producing RNA molecules. These RNA molecules are necessary in order to produce proteins.

1.2 Proteins

Proteins are biological molecules with a high variety of functions. Many known proteins contain around 100 amino acids but there are also bigger proteins that contain several hundreds of amino acids. The first group of proteins is the cytoskeleton. With regard to the structure of the cell it is the most important component, responsible for cell division and the establishment of structure. A second group consists of membrane proteins. Some of these proteins supply the cell with glucose and other molecules necessary for its nutrition while receptor molecules are necessary for the transduction of signals. A third group of proteins processes the glucose into chemical energy equivalents like **adenosine triphosphate** (ATP). While the information for

replication of the cell is stored on nucleic acids, processing and organization of the replication process is mainly realized by proteins. Yet another group of proteins are enzymes. Enzymes catalyze chemical reactions which would run only very slowly or not at all without their influence. A frequent principle in this connection is the stabilisation of unstable inter-medium-states of a reaction. Last but not least, another important function of proteins is the transduction of signals. The functionality of proteins is defined by their structure.

1.3 The structure of proteins

Proteins consist of α -amino acids that are connected to a chain. All amino acids contain an amine- ($-\text{NH}_2$) and a carboxylic- ($-\text{COOH}$) group which are bound to the C^α atom forming the backbone (shown in figure (2)).

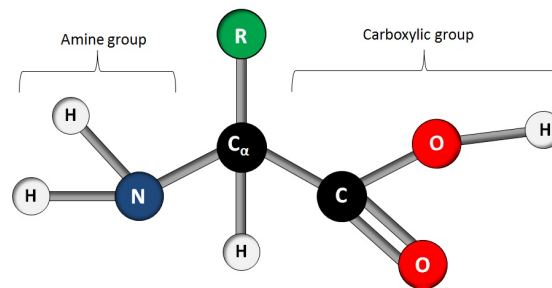


Figure 2: Basic structure of amino acids (adapted from: Branden and Tooze^[2])

Every amino acid contains a sidechain (R). In proteins, amino acids are connected by peptide-bonds. The peptide-bond has a double-bond character, that is why almost no rotation around the ω -angle occurs (shown in figure (3)). The structure of a protein is mainly defined by its conformation in ψ -angle and ϕ -angle. Figure (3) shows two tautomeric structures of amino acids in a protein.

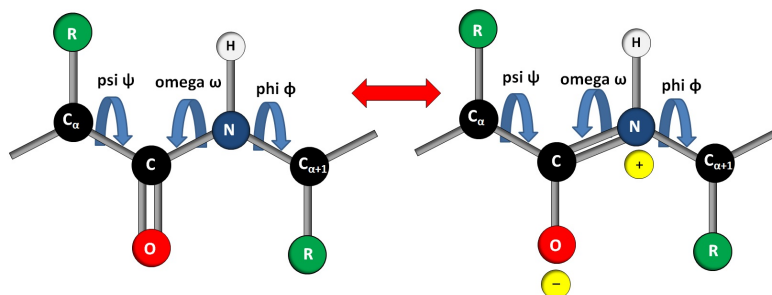


Figure 3: Peptide bond of amino acids. Blue arrows indicate bond angles (adapted from: Branden and Tooze^[2])

The functionality of a protein is determined by the 3D arrangement of its amino acid chain ([3]). The chemical - or primary - structure defines the 3D structure of the protein. However, the relationship between primary structure and 3D arrangement is not fully understood yet.

The first 3D structure of a protein was published in 1958 by John Kendrew, who determined the structure of myoglobin using x-ray diffraction (Kendrew *et al.*, 1958^[4]). Kendrew pointed out that "perhaps the most remarkable features of the molecule are its complexity and its lack of symmetry. The arrangement seems to be almost totally lacking the kind of regularities which one instinctively anticipates, and it is more complicated than has been predicted by any theory of protein structure" ([2]). Although the structure of proteins varies, there are some regularities in protein structure such as the appearance of secondary structure motifs, like α -helix and β -sheet (shown in figure 4), which are also called domains. These domains show typical compositions of amino acids and can be considered to be the building blocks of proteins.

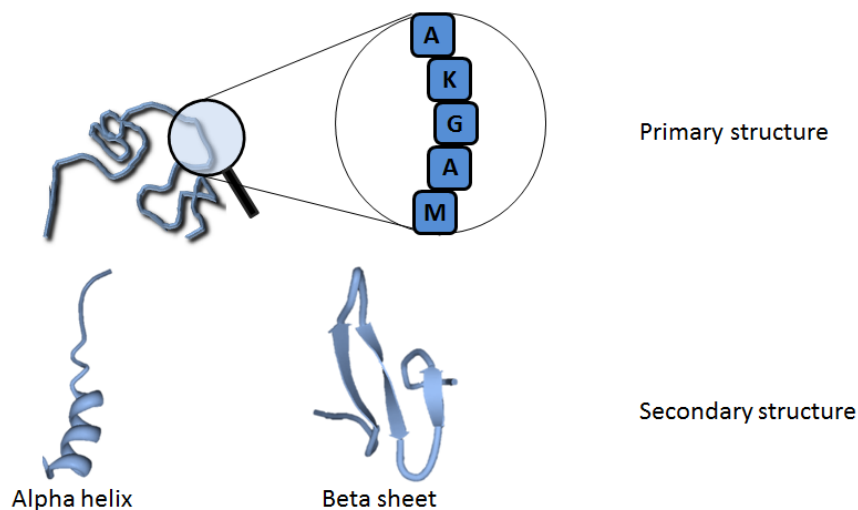


Figure 4: Primary and secondary structure of proteins (adapted from: Buchner and Kieferhaber ([3]))

Proteins usually consist of several domains that are connected by flexible linker regions. This combination of domains is called the tertiary structure of a protein (shown in figure 5). It is stabilized by non-covalent hydrophobic interactions as well as specific hydrogen bonds, salt bridges and disulfide bridges between specific amino acid side chains of the protein.

The formation of several proteins forming one protein complex is called quaternary structure (shown in figure 5). The single proteins of one complex are connected by the same non-covalent hydrophobic interactions, hydrogen bonds, salt bridges and disulfide bridges as the tertiary structure of a protein.

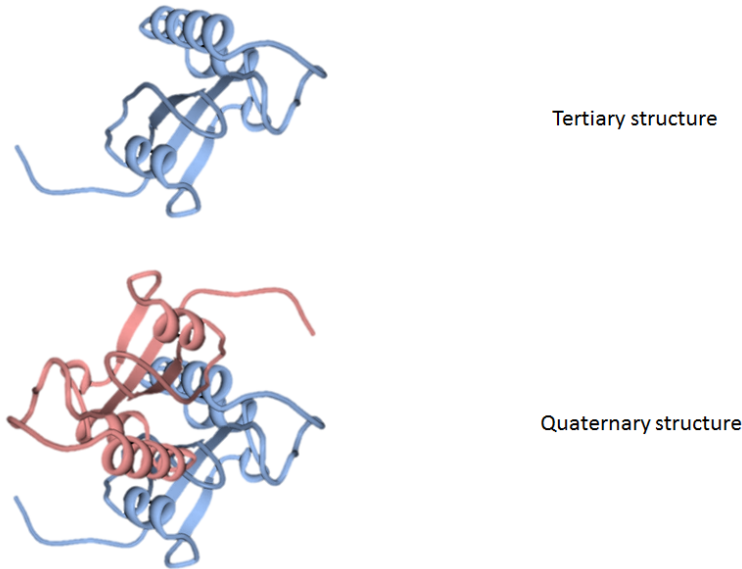


Figure 5: Tertiary and quaternary structure of proteins (adapted from: Buchner and Kieferhaber (^[3]))

1.4 Protein folding

A cell permanently produces new protein molecules. They are synthesized as a chain of amino acids that has no specific structure. Over time, the chain autonomously folds into its functional structure. Protein folding is the physical process in which a protein turns from a random coil formation into its native structure spontaneously. Information necessary for the native folding of the protein is given in its primary structure. In contrast, some proteins are not able to fold spontaneously on their own but rather need the help of chaperones to fold into their native structure. Yet another exception are membrane proteins that need a plasma membrane to reach their 3D structure.

In 1961 Anfinsen showed the reversible folding and unfolding of ribonuclease A (Anfinsen *et al.*, 1973^[5]). This protein contains eight cysteine residues that could theoretically form 105 different disulfide bonds to stabilize a certain protein structure. Anfinsen predicted that the driving force for the folding of proteins must be a free energy gradient. His theory has become known as the Anfinsen dogma, also

called the thermodynamic hypothesis. It predicts that the native structure has a minimum of Gibbs free energy. The Gibbs free energy ΔG is a thermodynamic potential that can be used to describe chemical reactions. The value of ΔG indicates whether a chemical reaction is either thermodynamically favorable and will occur spontaneously (exothermic) or if it is thermodynamically unfavorable and will not occur spontaneously (endothermic). Exothermic reactions with a negative ΔG value will emit heat while endothermic reactions need to absorb energy from the environment to take place. Equation (1) shows the dependence of Gibbs free energy ΔG :

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

In equation (1), ΔH is the enthalpy which describes the total energy of a system, T is the temperature of the system and ΔS is the entropy which is a measure of disorder for the system. A folding funnel is a simplified model related to the Anfinsen dogma. Like the Anfinsen dogma, the folding funnel shows the Gibbs free energy on the z axis and folding intermediates on the x axis (shown in figure (6)).

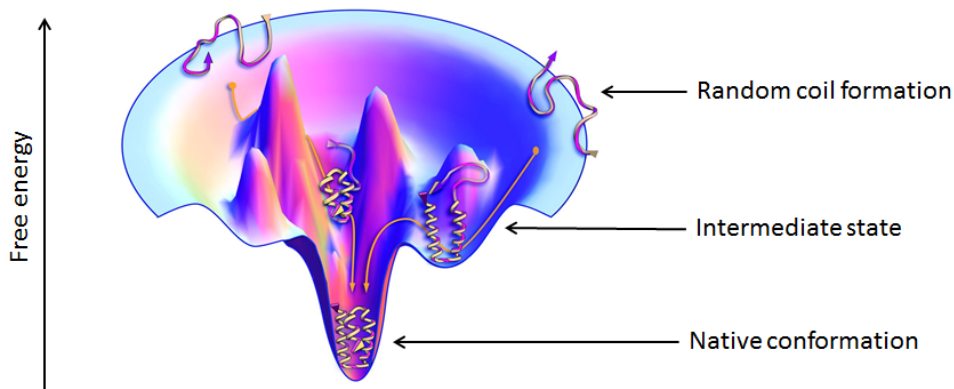


Figure 6: Schematic illustration of the protein folding funnel (taken from: ^[6])

Thus, in protein folding, several conditions have to be met according to the Anfinsen dogma. The first condition is the uniqueness of the minimal free energy state. Secondly, the protein has to be stable so as to guarantee that small changes in the surroundings do not change its folding state. The final condition is the kinetical accessibility of the intermediate states. The path in the folding funnel must be reasonably smooth so that it is not necessary for the protein to undergo high structural changes. Hence, the folding to the native state occurs spontaneously under these conditions of the Anfinsen dogma.

In 1965, John Brandts formulated the hypothesis that the thermodynamics of protein unfolding are dominated by hydrophobic free energy. He made his prediction based on the shape of unfolding curves and proposed that the difference in heat capacity ΔC_P between the native and the unfolded state, has a large positive value. Based on these ideas the theory of hydrophobic collapse was formulated. It predicts that hydrophobic free energy of the polypeptide chain is reduced by placing amino acids with hydrophobic sidechains in the less water-facing inner parts of the protein while amino acids with hydrophilic sidechains face the surrounding water at the outer parts. The hydrophobic collapse theory is closely related to the folding funnel theory in terms of free hydrophobic energy being the driving force for protein folding ([7]).

In 1969 Levinthal rationalized that protein folding would need to establish all possible backbone conformations randomly if it occurred as a two state reaction without intermediates. His calculations showed that the time necessary for a random search folding process might in some cases take longer than the lifetime of the universe. From that finding he concluded that there had to be intermediate states. Since the process usually only takes some seconds there has to be a bias in free energy between two intermediate states during protein folding. Levinthal formulated the hypothesis that if more than one pathway is available, protein folding will always follow the fastest one. In 1996, his hypothesis was confirmed experimentally for a protein from the serpin family that is able to build two different stable structures ([3]).

1.5 Beta sheets and WW-domains

An important domain structure in proteins is the β -sheet. The β -sheet consists of laterally connected β -strands. The connection of these β -strands is established by two or more hydrogen bonds between the C=O group of the C $^\delta$ atom and the N-H group of the nitrogen in the peptide backbone. The β -sheet forms a pleated, sheet-like, threedimensional structure.

Two different orientations of β -strand connections are known. One is the anti-parallel- β -sheet where the ends of the C-terminals of the two β -strands point into opposite directions (shown in the blue upper part of figure (7)).

The other one is the parallel- β -sheet where the C-terminal ends of the two connected β -strands point into the same direction (shown in the green lower part of figure (7)). β -sheets can consist of anti-parallel, parallel or a combination of both orientations of β -strand connections ([2]).

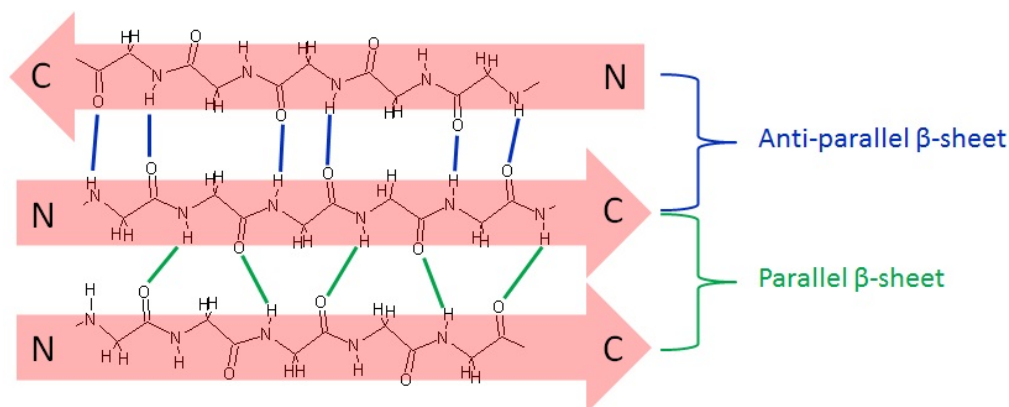


Figure 7: Schematic illustration of β -sheets in proteins (adapted from: Buchner and Kieferhaber ([3]))

The WW-domain is a widespread structural motif that can be found in a variety of unrelated proteins. Its name arises from the two conserved tryptophanes in its polypeptide chain (W is the one letter code for tryptophan (Staub *et al.*, 1996^[8])). Most WW-domains have a typical length of 36-40 amino acids. Their main biological function is protein-protein interaction by binding to proline-rich regions of target proteins. The WW-domain is an anti-parallel β -sheet that consists of three β -strands which are connected by two linker regions (shown in figure (8)).



Figure 8: Primary structure of the hPin1-WW-domain. Amino acids emphasized in red are inside the β -strands.

The specificity of its native folding and its structural stability are dominated by two types of interactions. On the one hand, there is the interaction of hydrophobic clustering. One hydrophobic cluster is mainly formed by the residues Arginine14, Tyrosine23 and Phenylalanine25 (shown in figure (9) A) (green)).

Another hydrophobic cluster lies on the opposite side of the β -sheet and is mainly formed by the residues Leucine7, Tryptophan11, Tyrosine24 and Proline37 (shown in figure (9) B) (red)).

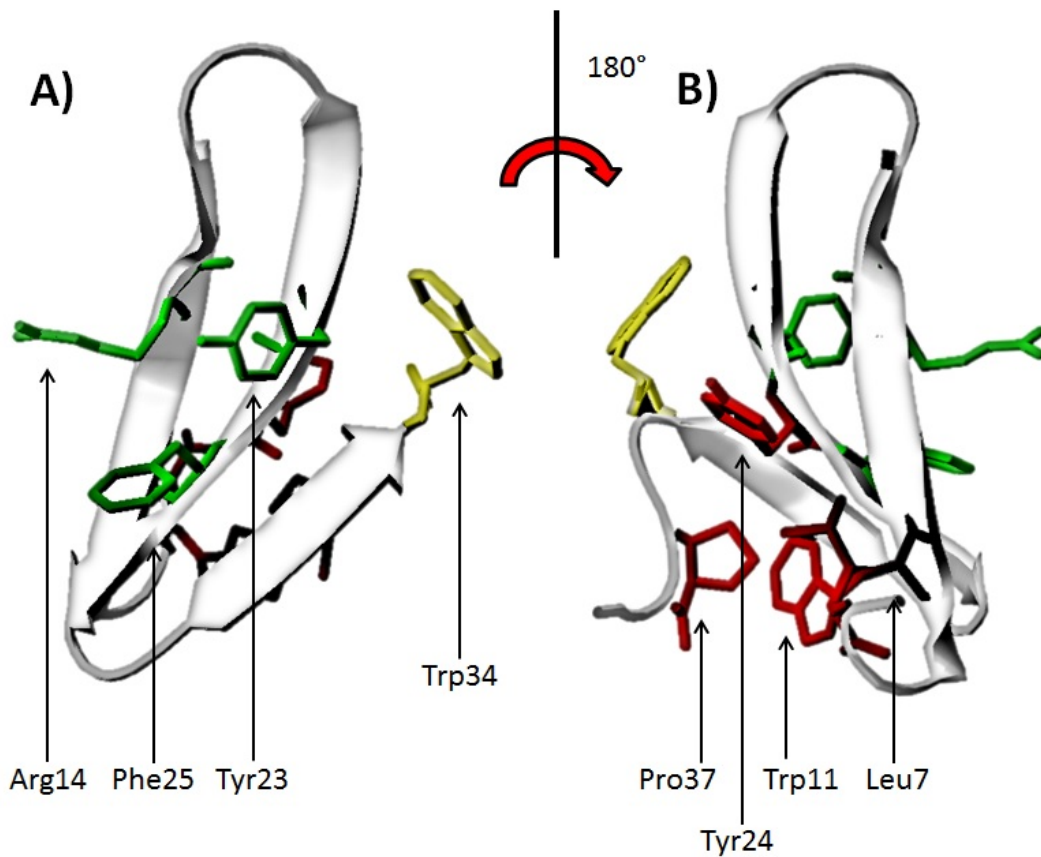


Figure 9: Hydrophobic clusters of hPin1-WW-domain. Amino acids in red and green form the hydrophobic clusters (adapted from: Protein data bank identification: Pin1 ([9]))

The second type of structure-stabilizing interaction, on the other hand, is a network of H-bonds and salt bridges formed between the backbones and sidechains of eight different amino acids. The Glutamine12 is connected to the Arginine14 by a salt-bridge and Histidine27 by H-bond (shown in figure (10, A)).

The Asparagine26 is connected by H-bonds with Proline9 and Tryptophan11 thereby connecting strands 1 and 2 with each other. Furthermore, Asparagine26 is also connected to Isoleucine28 and Threonine29 by H-bonds, (shown in figure (10, B)). A study on the stabilizing function of hydrophobic clusters and the H-bond network was published by Jäger *et al.* 2001 (Jäger *et al.*, 2001^[10]).

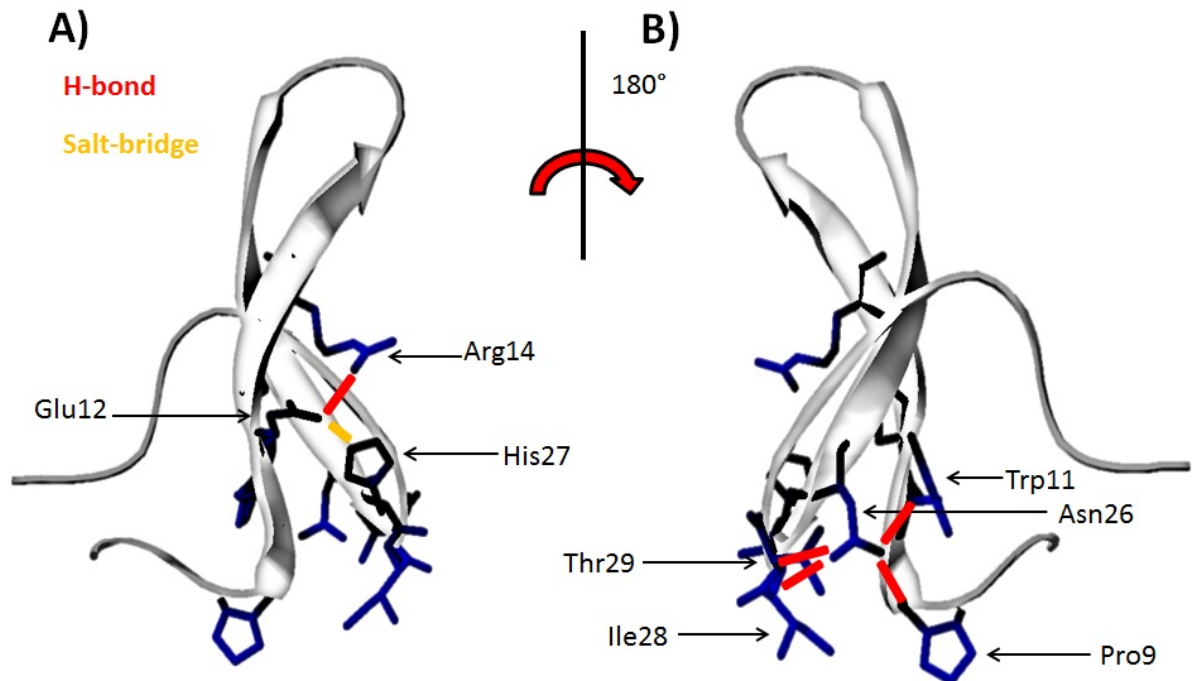


Figure 10: H-bond network of hPin1-WW-domain. H-bonds are shown in red while salt bridges are shown in yellow (adapted from: Protein data bank identification: Pin1 (^[9]))

1.6 Goal of this project

The goal of this PhD thesis was to analyze the folding properties of small peptide domains by single molecule FRET measurements. The hPin1-WW-domain was used as a model system to obtain information about the folding of small β -sheet proteins. The human rotamase (hPin1) protein is a mitotic regulator which is essential for the transition of the cell cycle.

Many diseases result from a loss of protein function due to a loss of native protein structure. The best known examples are the Creutzfeldt-Jakob disease and amyloidosis. In the case of amyloidosis, proteins lose their native structural properties and transform into long chains of polymerized β -sheets which precipitate and form long

fibrils called amyloids. For medical purposes, it would be very interesting to predict the structure and function of a protein from its primary sequence. Upon obtaining this information, it might be possible to design drugs and therapies against many diseases through the engineering of proteins (^[7]).

One feasible method to get information about the structure and function of proteins is the **single molecule** (sm) FRET. Using FRET, it is possible to measure the distance between two fluorophores. This being said, it becomes clear that it is impossible to get information about the entire structure of a protein from one FRET experiment. In most FRET studies concerning protein folding, the proteins are unfolded by chemicals like guanidinium hydrochloride or urea or thermally unfolded. The main advantage of smFRET experiments is that they have a high temporal resolution which can be used to analyze fast dynamics. FRET experiments can be used to analyze inter- and intramolecular dynamics like folding and unfolding properties of proteins.

The upper bound for protein dynamics had been determined by Chung *et al.* (2012^[11]) to be around 200 μ s.

2 Materials and Methods

2.1 Materials

All chemicals were obtained from Sigma-Aldrich (Hamburg, Germany), Merck (Darmstadt, Germany), Invitrogen (Hamburg, Germany), and Roth (Karlsruhe, Germany).

2.2 Plasmid DNA mini preparation

The mini preparation of DNA was accomplished with a plasmid mini kit (VWR, Darmstadt, Germany) according to the manufacturers protocol. Single bacterial colonies were picked and inoculated into 5 ml of LB media containing 100 µg/ml ampicillin. The bacterial cells of the strain BL21 Gold were grown in the incubator at 160 rpm for 16 h at 37 °C. Subsequently, these cells were collected by centrifugation at 5000 rpm for 15 min and resuspended in mini-prep buffer-I (50 mM Tris, 10 mM EDTA, pH 8.0). The cells were lysed by mini-prep buffer-II (0.1 M NaOH, 1% SDS). After the proteins in the solution had been precipitated by buffer-III (1.5 M potassium acetate, pH 5.5), the supernatant phase was transferred to DNA binding columns. Bound to the columns, the DNA was washed with mini-prep wash-buffer (70 % ethanol, 30 % H₂O) by centrifugation at 5000 rpm for 1 min. Finally, the DNA was eluted from the columns with 50 µl of H₂O by centrifugation at 5000 rpm for 1 min and stored at -20 °C.

2.2.1 Purification of DNA

Purification of plasmid DNA was carried out with a Zymo-clean kit (Zymo Research, Irvine, USA) according to the manufacturers protocol. The DNA was transferred to DNA binding columns by centrifugation at 5000 rpm for 1 min. The DNA bound to the columns was cleaned with wash-buffer by centrifugation at 5000 rpm for 1 min. DNA was eluted from the columns with 10 µl to 30 µl of H₂O by centrifugation at 5000 rpm for 1 min and afterwards the DNA was stored at -20 °C.

2.2.2 Determination of nucleic acid concentration

The concentration of nucleic acids was determined with a Nanodrop2000 (Wilmington, USA) microvolume spectrometer. For the analysis, 2 μl of DNA solution was used. The absorbance of the DNA was measured at 260 nm and the concentration of the DNA was calculated with the Lambert-Beer law (which is further explained in (3)).

2.2.3 Restriction of nucleic acids

The restriction of DNA was achieved by using restriction endonucleases (NEB, Frankfurt, Germany), i.e. enzymes specialized to cut DNA at specific palindromic sequences. There are about a hundred different restriction endonucleases available, each one specifically cutting one certain palindrome. A typical restriction mixture has a total volume of 20 μl - 50 μl and contains up to 10 % of restriction endonuclease. Usually, one to three units of restriction endonuclease were used for 1 μg of DNA. The respective restriction buffer which had been delivered by the supplier was added to the restriction mixture in order to achieve optimal restriction conditions. The restriction mixture was incubated for 3 h at 37 °C. Table (1) shows the mixture of a 20 μl restriction preparation.

DNA (100 ng/ μl)	4 μl
Restriction endonuclease	2 μl
Restriction buffer	2 μl
H ₂ O	12 μl

Table 1: Restriction mixture (20 μl)

2.2.4 Agarose gel electrophoresis

The DNA was separated for analysis by agarose gel electrophoresis. To prepare a 1 % agarose gel of 60 ml, 0.6 g of ultra pure agarose (Invitrogen, Hamburg, Germany) was dissolved in 60 ml of TAE buffer. The mixture was heated in a microwave, at 800 Watt for 2 - 4 min, until the solution became clear.

The melted agarose was cast into a gel chamber (Whatman (GE), Frankfurt, Germany) and kept at RT for 30 min until it was jellied (composition of the gel is shown in table (2)).

TRIS HCl	4.84 g
Acetic acid 100 %	1.14 g
EDTA 0.5 M pH 8.0	2 ml
H ₂ O	998 ml

Table 2: TAE buffer (1 l)

Afterwards, a loading dye (composition shown in table (3)) was added in order to prepare the DNA samples for the agarose electrophoresis. To analyze the length of the DNA fragments a 1 kb DNA standard (Fermentas, Darmstadt, Germany) was used. Following this, the electrophoresis was run in a electrophoresis chamber (Biorad, München, Germany) applying 110 Volt for 60 min.

Bromophenol blue 1 %	5 ml
Glycerole	3 ml
H ₂ O	2 ml

Table 3: DNA loading dye (10 ml)

2.2.5 DNA extraction from agarose gels

DNA fragments separated in agarose gels were extracted using a zymo-extraction kit (Zymo Research, Irvine, USA). The DNA fragments of interest were cut out of the agarose gel and mixed with 300 μ l agarose dissolving buffer (ADB). Subsequently, the mixture was dissolved through incubation at 50 °C for 10 min. In a next step, the dissolved agarose was loaded on a Zymoclean DNA column by centrifugation at 10000 rpm for 1 min. Bound DNA was washed with 200 μ l wash-buffer (70 % ethanol, 30 % H₂O). Residual wash-buffer was removed by centrifugation at 10000 rpm for 1 min. Finally, 20 μ l H₂O were added to the column and centrifuged at 10000 rpm for 1 min to elute the DNA from the column.

2.2.6 Ligation of DNA

The ligation of plasmid vectors and DNA fragments (inserts) was performed by using T4 DNA ligase (NEB, Frankfurt, Germany). The insert-DNA was added in a molar concentration three times higher than that of the plasmid vector DNA. The respective ligation buffer was added to the ligation mixture in order to achieve optimal ligation conditions. Ultimately, the ligation mixture was incubated for 12 h at 10 °C.

2.3 Transformation of nucleic acids

The DNA of the expression vector was transformed into electrocompetent cells of the bacterial strain *E. coli* BL21 Gold by electroporation. A 50 µl aliquot of competent bacteria was charged with 1 µl DNA from a 100 ng/µl stock. Then, an electroporation cuvette with a 1 mm gap was used to expose the mixture to a 3 ms - 5 ms pulse of 1800 Volt in a micropulser (Biorad, München, Germany). Subsequently, the bacteria were transferred into 200 µl LB media and stirred at 600 rpm for 1 h at 37 °C. Afterwards, the grown bacteria were plated on an agarose plate containing 100 µg/ml antibiotics (either ampicillin or kanamycin depending on the resistance gene of the expression vectors) and grown for 16 h at 37 °C.

2.4 Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) is the standard method to amplify DNA fragments (Mullis *et al.*, 1986^[12]). In several cycles, the template DNA is melted and its strands are separated. In the following annealing step, short oligo nucleotides (primers) are hybridized to the template strands upon cooling. In the elongation phase, a polymerase enzyme is used to elongate the new DNA molecule. After the amplification is finished, the new DNA fragment can be restricted and ligated into a DNA vector as explained before.

2.4.1 QC Polymerase chain reaction

A Quickchange polymerase chain reaction (QC-PCR) was used to insert point mutations into DNA which will lead to a mutation in the amino acid sequence of the corresponding protein. In QC-PCR, the concentration of both primer and vector DNA has to be determined accurately. Thus, 125 ng of the reverse primer and forward primer had to be added into a 50 μl reaction, whereas the concentration of the vector DNA had to be 25 ng. A 50 μl reaction contained 5 μl 10x Pfu polymerase buffer, 1.25 μl Primer fwd (100 ng/ μl stock), 1.25 μl Primer rev (100 ng/ μl stock), 0.2 mM of each dNTP, 2.5 μl template DNA (10 ng/ μl stock), and 2.5 U Pfu polymerase. The amplification was carried out in a thermocycler using the cycle shown in table (4). PCR steps 2 to 4 were repeated 12 times.

<i>Step</i>	<i>Temperature</i> [$^{\circ}\text{C}$]	<i>Time</i> [min]
1 initial denaturation	95	0:30
2 cycle denaturation	95	0:30
3 annealing	55	1:00
4 elongation	68	1:00
5 final elongation	68	14:00

Table 4: QC PCR cycle

2.4.2 Primer design for QC PCR

QC primers contain a mismatching triplet to insert a point mutation into the plasmid. Primer regions matching the template DNA surrounded the region of mismatch. The primers were calculated and designed in order to have the correct T_m , using the following equation provided by Agilent technologies (2):

$$T_m = 81.5 + 0.41(\% GC) - (675/N) \quad (2)$$

T_m is the melting temperature at which the primer will be separated from the DNA template. This melting temperature depends mainly on the GC (guanine and cytosine) content of the primer. This GC content is the ratio of GC pairs which form three hydrogen bonds, in contrast to AT (adenine and thymine) pairs that form two hydrogen bonds. Due to their higher thermo-stability, GC pairs increase the T_m of a primer.

The primers used in this work had a length between 25 and 45 bases and a melting temperature (T_m) of ≥ 78 °C. In the equation above, N stands for the number of bases in the primer, not counting the bases being inserted.

2.4.3 DpnI digestion

Maternal template DNA was removed from the PCR reaction by a digestion with DpnI endonuclease. The DpnI endonuclease specifically digests methylated DNA. Therefore, it is possible to remove the methylated maternal template DNA from the PCR reaction while the amplified DNA, which is non-methylated, is kept. The DpnI digestion was performed with 20 U of DpnI over 12 h at 10 °C.

2.5 *E. coli* cell culture

2.5.1 Bacterial strains and DNA plasmid vectors

In this thesis the following bacterial strains and DNA plasmid vectors were used:

Bacterial strains:

BL21 Gold (Agilent Technologies, Böblingen, Germany): cells of this strain are a T7 RNA polymerase-based expression system which is designed for high-level protein expression.

OverExpress(tm)C41(DE3) (Lucigen, Middleton, USA): cells of this strain are highly tolerant against toxic proteins.

OverExpress(tm)C41(DE3)plysS (Lucigen, Middleton, USA): cells of this strain are highly tolerant against toxic proteins. The cells contain pLysS plasmid which encodes T7 phage lysozyme which, in turn, is inhibitory for T7 polymerase. Therefore, it reduces and eliminates almost all expressions from plasmids containing T7 promoter when not induced by IPTG.

OverExpress(tm)C43(DE3) (Lucigen, Middleton, USA): cells of this strain are highly tolerant against toxic proteins.

OverExpress(tm)C43(DE3)plysS (Lucigen, Middleton, USA): cells of this strain are highly tolerant against toxic proteins. The cells contain pLysS plasmid which encodes T7 phage lysozyme which, in turn, is inhibitory for T7 polymerase. Therefore, it reduces and eliminates almost all expression from plasmids containing T7 promoter when not induced by IPTG.

XL1-Blue (Stratagene, Hamburg, Germany): Cells of this strain are tetracycline resistant.

DNA plasmid vectors:

pET11a (Novagen, Madison, USA): contains a T7 promotor for induction of expression and a kanamycin resistance gene for selective growth (vectormap shown in (55)).

pET24b (Novagen, Madison, USA): contains a T7 promotor for induction of expression and a kanamycin resistance gene for selective growth (vectormap shown in (56)).

pET27b (Novagen, Madison, USA) : contains a T7 promotor for induction of expression and an ampicillin resistance gene for selective growth (vectormap shown in (57)).

pGEX2T (Novagen, Madison, USA): expressed proteins will be fused to **glutathione-S-transferase** (GST). The linker between GST and expressed proteins contains thrombin recognition sites for cleavage. Contains T7 promotor and ampicillin resistance gene (vector shown in figure (58)).

2.5.2 LB media

The lysogeny broth (LB) media is used for bacterial cell culture. The LB composition is listed in table (5):

Trypton 1 %	10 g
Yeast extract 0.5 %	5 g
Sodium chloride 0.5 %	5 g
H ₂ O	1 L

Table 5: LB media

2.5.3 Preparation of competent cells

Bacterial cells of the strain *E. coli* BL21 Gold were grown in a 6 ml LB pre-culture in the incubator at 160 rpm for 16 h at 37 °C. The pre-culture was added to 350 ml LB/antibiotics and grown at 160 rpm at 37 °C until it reached an optical density (OD_{600}) of 0.6 - 0.8 at 600 nm. Subsequently, the bacterial cells were collected by

centrifugation in a Beckman Coulter Avanti J26 xp centrifuge using a Js 5.3 rotor at 5000 rpm for 15 min at 4 °C. After the supernatant had been discarded, cells were resuspended in 300 ml buffer (ice-cold water containing 10 % of glycerol). Following the repetition of the centrifugation step, the pellet was resuspended in 20 ml buffer. Ultimately, after a final centrifugation step, the cells were resuspended in 2 ml buffer, separated into 50 µl aliquots and stored at -80 °C.

2.5.4 Protein Expression

Protein expression was performed using bacterial *E. coli* cells which had been transformed using the expression vector. Cells transformed with the respective expression vector were grown in a 50 ml LB/ antibiotics pre-culture in the incubator at 160 rpm for 16 h at 37 °C. For expression, 4 liters of LB/ antibiotics were inoculated with the pre-culture and grown at 37 °C. Once it had reached an optical density of 0.4 - 0.6 at 600 nm (OD_{600}), Isopropyl- β -D-thiogalactopyranosid (IPTG) was added to induce the expression of the protein of interest. The expression ran at 160 rpm for 16 h at 22 °C. Afterwards, the bacterial cells were collected by centrifugation at 5000 rpm for 15 min and stored at -80 °C.

2.6 Purification of recombinant proteins

2.6.1 Protein purification by nickel-NTA

The bacterial cells were lysed and the proteins were purified via a poly **H**istidine affinity tag (His-tag). To do so, a pellet from a 500 ml expression culture was resuspended in 10 ml PBS (pH 7.4) containing lysozyme (200 µg/ml) in order to induce the lysis of the cells. Subsequently, the solution was sonified using a Sonoplus sonifier with a MS73 Sonotrode (Bandelin, Berlin, Germany) for 6 min while the bacterial solution was cooled on ice. The lysate was transferred into ultracentrifugation tubes and centrifuged in a Beckman Coulter Optima L 90 K Ultracentrifuge using a Ti70 rotor at 40000 rpm for 30 min at 4 °C. In the following, the supernatant was transferred to a column containing a nickel-NTA matrix (Macherey-Nagel, Düren, Germany) with a bed volume of 5 ml, which had been washed and equilibrated with 50 ml PBS (pH 7.4). Afterwards, the supernatant containing the His-tagged protein was incubated on this nickel-NTA column for 1 h on ice. The flow-through which contained unbound proteins and other cellular components was discarded.

50 ml PBS (pH 7.4) were added to wash non-specific bound proteins from the column. 20 ml PBS (pH 7.4), containing 10 mM imidazole, were added to wash weakly bound proteins without His-tag from the column. The His-tagged protein was eluted with 10 ml elution buffer (PBS pH 7.4, 300 mM imidazole) and stored at -80 °C.

2.6.2 Protein purification by glutathione sepharose

The bacterial cells were lysed and the proteins purified using a glutathione S-transferase affinity tag (GST-tag) as follows: A pellet from a 500 ml expression culture was resuspended in 10 ml glutathione sepharose binding buffer (PBS 140 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.3). In order to induce the lysis of the cells, Lysozyme (200 µg/ml) was added. Subsequently, the solution was sonified using a Sonoplus sonifier with a MS73 Sonotrode (Bandelin, Berlin, Germany) for 6 min while the bacterial solution was cooled on ice. The lysate was transferred into ultracentrifugation tubes and centrifuged in a Beckman Coulter Optima L 90 K Ultracentrifuge using a Ti70 rotor at 40000 rpm for 30 min at 4 °C. The supernatant was transferred to a glutathione sepharose column with a bed volume of 5 ml, which had been washed and equilibrated with 50 ml glutathione sepharose binding buffer to assure best binding affinity for the GST-tag. Afterwards, the supernatant containing the protein of interest, which was linked to the GST (from now called: protein-GST), was incubated on this glutathione sepharose column for 1 h on ice. The flow-through which contained unbound proteins and other cellular components was discarded. 50 ml Glutathione sepharose binding buffer were added to wash non-specific bound proteins other than the protein-GST from the column. The protein-GST was eluted by 10 ml elution buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0) and stored at -80 °C.

2.6.3 Thrombin proteolysis

Thrombin protease was used to cleave the protein from the GST by proteolytic cleavage at the thrombin recognition site. First, the protein-GST had to be rebuffed in a PD10 desalting column in order to preserve optimal conditions for the enzymatic cleavage. The Sephadex-25 matrix of the PD10 column had a cutoff size of 5000 Da. The column was used to exchange the glutathione containing buffer which surrounded the protein-GST to PBS. After 1 ml of the protein-GST had been loaded on the column, it was eluted with PBS pH 7.3. The first 1 to 1.5 ml of the elution

contained the protein-GST in PBS buffer without glutathione. 500 U of thrombin (GE, Frankfurt, Germany) were used to cut 8 mg of protein-GST. Cleavage was performed for 12 - 16 hours at RT. Afterwards the mixture was stored at -80 °C.

2.6.4 Size exclusion chromatography by HPLC

The free protein was separated from the GST-tag, uncleaved protein-GST and thrombin protease by size exclusion chromatography using high performance liquid chromatography (HPLC). For this procedure, a Jasco MD2010 HPLC with a Superdex Peptide (GE, Frankfurt, Germany) column was used. The separation range of the column was 100 - 7000 Da. Bigger molecules such as uncleaved protein-GST (26+x kDa), GST (26 kDa) and thrombin (36 kDa) were excluded due to their size by the columns void-volume of 20 kDa. For the separation, 1 ml of cleavage mixture was loaded on the column and the elution was done by PBS pH 7.4 with a flow rate of 0.4 ml/min.

2.7 Determination of protein concentration

For the determination of the protein concentration, the absorbance of the protein at 280 nm was measured and the molar extinction coefficient ϵ of the protein was calculated. So, the absorbance of the hPin1-WW-domain was measured in a spectral photometer (Jasco, Jena, Germany) and calculated by the Lambert-Beer law shown in equation (3):

$$\text{Absorbance} = \lg \frac{I_0}{I} = \epsilon \cdot c \cdot l \quad (3)$$

In equation (3) I_0 is the light intensity before the sample, I is the light intensity after the sample, c is the concentration of the sample [M], and l is the cuvette length [cm].

The molar extinction coefficient had been determined from the amino acid sequence of the protein (Gill *et al.*, 1998^[13]). For this, an algorithm provided on the internet platform ExPASy (ProtParam documentation^[14]) was used.

The following ϵ values of amino acids were used to calculate the ϵ values of the proteins (shown in table 6):

Amino acid	ϵ_{280}	number of amino acids
Tyrosine	1490	2
Tryptophan	5500	1
Cysteine	125	2

Table 6: Extinction coefficients of amino acids at 280 nm

The calculated molar extinction coefficient of the hPin1-WW-domain is 8730 [$M^{-1}cm^{-1}$]. The concentration of the hPin1-WW-domain was calculated from the absorbance at 280 nm divided by the molar extinction coefficient as shown in equation (4):

$$\text{Prot. conc. [M]} = \frac{A(280nm)}{\epsilon(\text{Protein})[M^{-1}cm^{-1}] \cdot l} \quad (4)$$

2.8 Electrophoresis

2.8.1 SDS-PAGE

Proteins were analyzed by 2D-denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The SDS-PAGE was conducted using an electrophoresis chamber (Biorad, München, Germany). Protein separation was performed in a polymer gel with 20 % polyacrylamide (resolving gel) and was located in the lower part of the SDS gel (mixture shown in table 7). This mixture was jellied for 20 minutes.

H ₂ O	2.45 ml
TRIS (1.5 M)	2.5 ml
Polyacrylamide	5.4 ml
SDS (20 %)	50 μ l
Ammonium persulfate	100 μ l
Tetramethylethylenediamine (Temed)	5 μ l

Table 7: Resolving gel 20 %

The upper part of the SDS gel, called stacking gel, had a polyacrylamide concentration of 4 % (mixture shown in table 8). In order to generate pockets for sample loading within the stacking gel, a comb was added to the upper part of the gel during polymerization. All in all, the SDS gel was ready for use after it had been jellied for 20 minutes.

H ₂ O	6.6 ml
TRIS (1 M)	1.5 ml
Polyacrylamide	1.33 ml
SDS (20 %)	50 µl
APS	50 µl
Temed	10 µl

Table 8: Stacking gel 4 %

For SDS-PAGE, the jellied gel was mounted into the electrophoresis chamber and filled with SDS buffer (mixture shown in table 9).

TRIS HCl	1.5 g
Glycin	7.2 g
SDS	0.3 g
H ₂ O	ad 500 ml

Table 9: SDS loading buffer (500 ml)

In the following steps, protein samples were mixed with SDS-loading dye (composition shown in table 10) and boiled at 95 °C for 10 min. The denatured protein samples as well as a molecular standard marker (Fermentas, Darmstadt, Germany) were added to the loading pockets of the stacking gel. The electrophoresis was performed applying 180 volt for 80 min.

β -Mercaptoethanol	500 μ l
SDS (20 %), 1.5 M	1 ml
Tris-HCl, 1 M, pH 6.8	1.25 ml
Bromphenolblue (20 %)	1.5 ml
Glycerol	1.5 ml
H ₂ O	19.25 ml

Table 10: SDS loading dye

2.8.2 Staining of SDS-PAGE

To stain the proteins which had been separated in the SDS-PAGE, coomassie brilliant blue R250 (Merck, Darmstadt, Germany) was used. The gel was incubated in 20 ml staining solution (mixture shown in table 11) and shaken for 1 to 2 h.

Coomassie brilliant blue R250 in H ₂ O (0.05 %)	50 ml
Acetic acid	20 ml
H ₂ O	30 ml

Table 11: Staining solution (100 ml)

After staining, the gel was transferred into 20 ml of destaining solution (mixture shown in table 12) to remove residual Coomassie brilliant blue from the gel. It was shaken for 60 min. Afterwards, a picture of the gel was taken in a gel documenter (UVIdoc, Biometra, Göttingen, Germany).

Methanol 100 %	40 ml
Acetic acid	20 ml
H ₂ O	50 ml

Table 12: Destaining solution (100 ml)

2.9 Mass spectrometry of proteins

The identification of purified GST-hPin1-WW-domain by **Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF)** mass spectrometry was carried out by Dr. Oliver Valerius from the Department of Microbiology at the University of Göttingen. MALDI samples of proteins were generated by cutting the corresponding bands out of the SDS gel. Further processing and the analysis itself were conducted by Dr. Oliver Valerius according to (Shevchenko *et al.*, 1996^[15]).

2.10 Nano-containers

The concept was to encapsulate proteins into polymerosomes which were used as nano-containers. Polymerosomes are spheres with a diameter between 50 nm and 200 nm which are assembled from polymers containing hydrophilic and hydrophobic parts. The nano-containers were immobilized by a streptavidin-biotin binding on a glass surface. Figure (11) shows a scheme of surface binding.

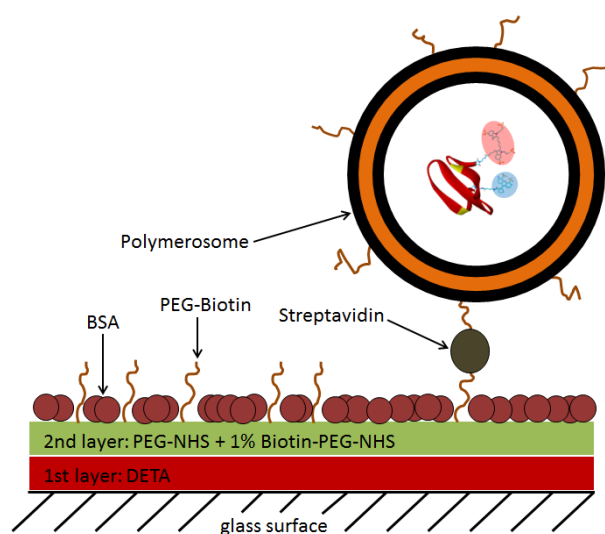


Figure 11: Scheme of nano-containers attached to the surface.

The polymerosomes were formed by amphiphilic ABA tribloc co-polymers of the type poly(2-methyloxazoline)-bloc-poly(dimethylsiloxane)-bloc-poly(2-methyloxazoline) = PMOXA-PDMS-PMOXA. Figure (12) shows a schematic drawing of the ABA tribloc co-polymer. The polymer was provided by the group of Wolfgang Meier (University Basel, Switzerland).

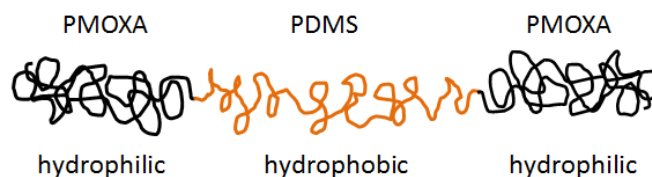


Figure 12: Schematic drawing of ABA tribloc co-polymer

The polymer was stored in chloroform at $-20\text{ }^{\circ}\text{C}$ at a stock concentration of 30 mg/ml for ABA and 0.5 mg/ml for ABA-biotin. For formation of polymerosomes, biotinylated and non-biotinylated ABA tribloc co-polymer were combined in a 2 ml glass tube. The mixing ratio of biotinylated to non-biotinylated polymer was 1:10 in terms of weight. Subsequently, chloroform was evaporated under permanent rotation at 100 rpm. One ml of PBS buffer containing the hPin1-WW-domain at a concentration of 100 pM was added drop by drop to the polymer film, followed by rotation at 100 rpm for 30 minutes. Due to this procedure, the polymeres formed multilamellar polymerosomes with inhomogeneous size. These polymerosomes were homogenized by extrusion trough membrane pores. Therefore, the solution was transferred to a Hamilton syringe and mounted into the extruder. The extrusion was done using a self-made automatic extruder that repeatedly pushed the solution through a Track-etch membrane (Whatman (GE), Frankfurt, Germany) with a pore size of 100 nm. After 400 cycles, the extrusion resulted in a solution of unilamellar polymerosomes. Immobilization of polymerosomes on a cover glass surface was achieved by streptavidin-biotin binding. 24 mm x 40 mm cover glasses were cleaned by sonication in a solution of 1 M potassium hydroxide (KOH) for 10 min, and then rinsed with water. The cleaned cover glasses were coated with 10 weight-% diethylentriamin (DETA) dissolved in water for 5 minutes. Subsequently, the cover glasses were dried in an oven for 30 minutes at $110\text{ }^{\circ}\text{C}$. 60 mg of PEG-NHS were dissolved in 60 μl Biotin-PEG-NHS of a 10 mg/ml solution for the coating of the glass surface. The mixture was added to 600 μl of cold aqueous KBO_3 (20 mM). A drop of 60 μl was placed at the center of a cover glass and another cover glass was placed on its top to be incubated for 1 h. For the purpose of preparing the chamber for the measurement, stripes of adhesive tape were placed on the cover glasses. An uncoated 22 mm x 22 mm cover glass was placed on top of the adhesive tape and gently pushed to adhere on the tape (construction scheme shown in figure 13).

Streptavidin dissolved in PBS at (pH 7.4) at a concentration of 10 ng/ μl was filled into the chamber and incubated for 30 min at RT. Afterwards, the streptavidin

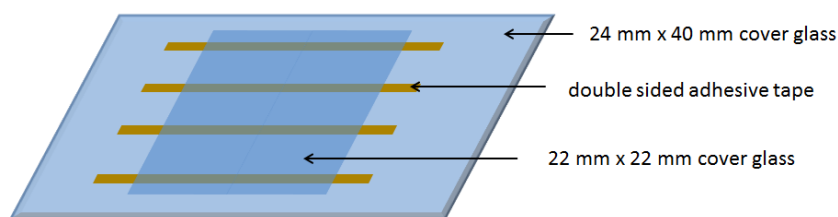


Figure 13: Chamber preparation for polymerosome measurement.

buffer was exchanged with 100 μM bovine serum albumine (BSA), dissolved in PBS (pH 7.4), to prevent the polymerosomes from touching the coated surface. In a next step, 20 μl of the polymerosome solution were added and was incubated for 30 - 60 minutes. Finally, PBS was used to wash away free polymerosomes from the channel. The resulting sample was measured using a microscope.

2.11 Circular dichroism (CD)

2.11.1 Chiral molecules

Chiral molecules are not congruent with their mirror image. An example of such a molecule is shown in figure (14). All α -amino-acids, except glycine, are chiral molecules. The chirality of amino acids arises from the asymmetric conjugation of the C_α atom. What is more, all chiral amino acids can exist in two different, mirror-symmetrical isomers that are called enantiomers. One isomer is called the D (dextrorotatory) and the other one is called the L (levorotatory). In nature, only L isomers of α -amino acids are found in proteins.

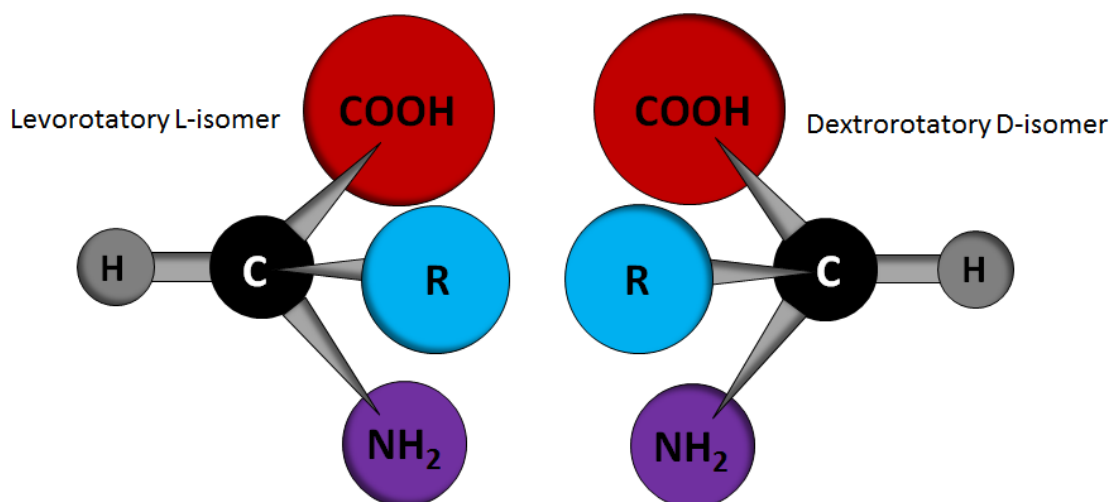


Figure 14: D and L isomers of aminoacids

2.11.2 Theory of circular dichroism

Circular dichroism (CD) is a common method to investigate the secondary structure of proteins (Greenfield *et al.*, 2007^[16]). Ultraviolet (UV) CD spectrometry uses the spectral range between 190 nm and 250 nm. Electromagnetic waves are jointly propagating oscillations of an electric and of a magnetic field. The field vectors oscillate perpendicularly to each other and to the propagation direction of the wave. If light is linearly polarized, the electric field oscillates in a single plane. This electric field vector can be considered as the sum of two basic polarization vectors.

One vector rotates clockwise E_R , while the other one rotates counterclockwise E_L . When linearly polarized light interacts with chiral molecules, E_R and E_L are absorbed to different extents.

After the light has passed the sample, the amplitude of E_R and E_L will be different, which results in an elliptical polarization (as shown in figure 15).

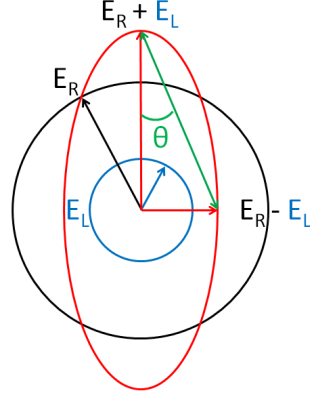


Figure 15: Circular dichroism. E_L is shown in blue, E_R is shown in black, the elliptical polarization is shown in red and angle theta is shown in green.

$$I_R = \frac{c}{8\pi} E_R^2 \rightarrow A_R = \log \frac{I_{R0}}{I_R} \quad (5)$$

The ellipticity of the resulting beam can be determined by measuring the absorbance of a right circular polarized beam (A_R) and of a left circular polarized beam (A_L). The difference in absorbance can be converted into the molar circular dichroism $\Delta \epsilon$ [$M^{-1} \text{cm}^{-1}$], shown in equation (6) ([3]).

$$\Delta \epsilon = \frac{A_R - A_L}{c \cdot l} \quad (6)$$

In equation (6) c is the concentration of the sample [M] and l is the cuvette length [cm]. The molar ellipticity θ (in degrees) is obtained by ($\Delta \epsilon \ll 1$).

$$\Delta \epsilon \left(\frac{\ln 10}{4} \right) \left(\frac{180}{\pi} \right) = \theta \quad (7)$$

In CD one is usually interested in the degree of molar ellipticity θ in the units [$\text{deg} \cdot M^{-1} \text{dm}^{-1}$] which is calculated by following equation:

$$[\theta] = 100\theta \approx 3298.2 \Delta \epsilon \quad (8)$$

2.11.3 Circular dichroism measurement

A Jasco J-815 CD spectrometer was purged with constant Nitrogen at a flow of 2.5 l/min during the entire measurement. This was necessary to protect optical elements from ozone which would arise if oxygen was irradiated by UV light. The samples were measured in Hellma quartz suprasil cuvettes with a path length of 10 mm. A range of 190 nm - 250 nm and a scanning speed of 20 nm/min were used for the UV-CD measurements. Measurements were repeated five times to achieve a satisfactory signal-to-noise ratio. Background absorption was determined by a blank measurement of PBS buffer. The concentration of the peptide in PBS buffer was between 25 μM and 50 μM . Subtraction of the blank measurement from the peptide measurement is done automatically by the Jasco software. The bands in a CD spectra arise from different conformations of the amide bonds in peptides and can be attributed to the presence of typical domain structures (Sreerama *et al.*, 2004^[17]). Thermal unfolding of peptides was performed by a Jasco ETC-273T temperature element. Measurements of maximum ellipticity at 226 nm were used for monitoring of the folding state of the hPin1-WW-domain. Figure (16) shows an example of CD data from hPin1-WW-domain.

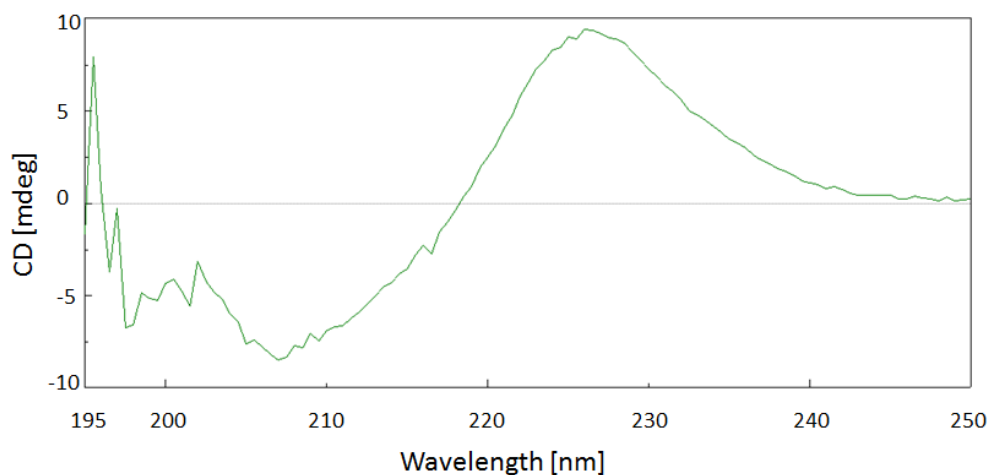


Figure 16: CD measurement evaluation. Shown is the measurement of the hPin1-WW-domain in PBS.

In this work the CD spectra were used to monitor the maximum ellipticity stemming from a loss of secondary structure due to unfolding of the protein.

2.12 Differential scanning calorimetry (DSC)

Differential scanning calorimetry (DSC) is a thermoanalytical method that is used to measure the heat capacity C_P of a sample. The heat capacity C_P is the amount of heat which is required to increase the temperature of a sample by a certain amount. In DSC measurements, the temperature is increased linearly while measuring the heat flow necessary to increase the temperature. The measurement is based on the difference in energy absorbance between the sample and the reference. With this referential measurement, it is possible to observe phase transitions such as crystallization, or melting of peptides ([3]). The following equations (9), (10) and (11) can be used to calculate the C_P value of a certain molecule.

$$\text{heat flow} = \frac{\text{heat}}{\text{time}} = \frac{\Delta E_{th}}{t} \quad (9)$$

$$\text{heating rate} = \frac{\text{temperature increase}}{\text{time}} = \frac{\Delta T}{t} \quad (10)$$

$$\text{heat capacity } C_P = \frac{\Delta E_{th}}{\Delta T} \left[\frac{J}{K} \right] \quad (11)$$

Rise in temperature leads to an unfolding of the protein. During unfolding, hydrogen bonds stabilizing the secondary structure of the peptide are broken. Thereby, energy is absorbed. This way, the peptide sample will absorb more heat than the reference buffer. Accordingly, the measurement showed that unfolding of the peptide was an endothermic process in which heat was absorbed from the environment (O'Neill *et al.*, 1964^[18]). Heat flow is the amount of heat that is taken up over time, the heating rate describes the change in temperature over time.

Applying DSC made it possible to measure the C_{mol} values of peptides. A pressure of 2.75 bar was used for DSC to ensure a constant concentration of the sample during the measurement. The total volume of the sample was 250 μl .

The experimental setup (VP-DSC, GE, Frankfurt, Germany) consisted of two measurement chambers (shown in figure 17).

Of these chambers the first one, called the sample chamber, was filled with the peptide sample. The second one (reference) contained the same buffer as the first but without any peptides being dissolved in it.

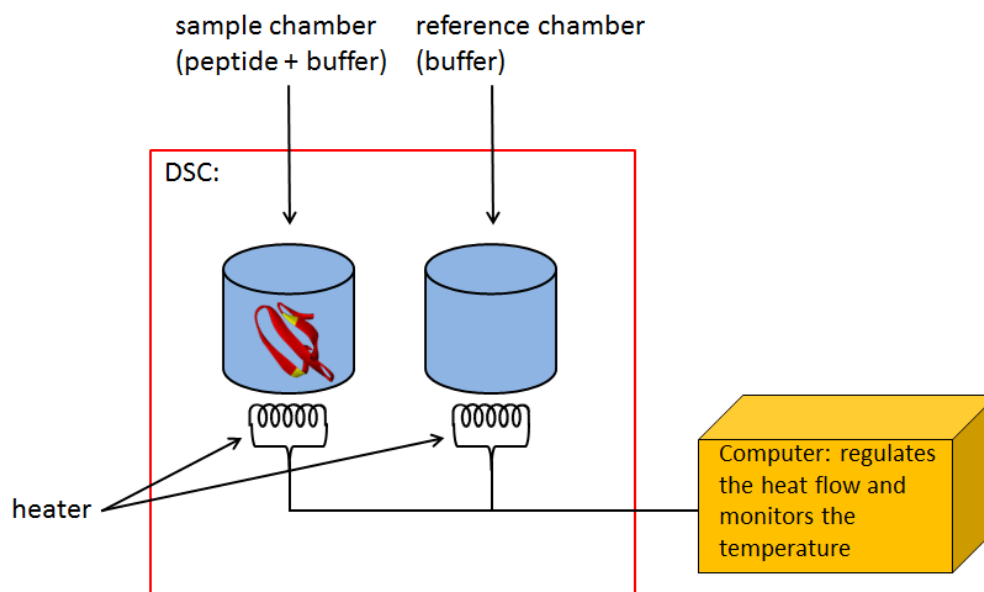


Figure 17: Schematic illustration of the DSC setup. Measurement chambers are shown in blue

Both, sample and buffer, had to be degassed prior to the measurement to avoid the appearance of air bubbles. Each chamber was surrounded by a heater. This was done in order to determine and subtract the background caused by the buffer. Thus, only the heat capacity of the protein could be observed. Afterwards, a measurement of buffer against buffer was performed to correct the baseline, a step necessary to accurately determine the heat capacity. In both measurements, the temperature of the chamber was increased linearly while at the same time determining the value of heat flow.

The required peptide concentration for a successful measurement is mainly dependent on the peptide size and its unfolding properties. For the measurement of the hPin1-WW-domain, a concentration of 600 μM was necessary to obtain a DSC signal. The heating steps during the measurement were 100 K/h at a pressure of 2.75 bar. The temperature range of the scan was 298 K (25 °C) to 368 K (95 °C).

2.13 FRET

2.13.1 Fluorescence

Today, fluorescence is a widely used effect in biochemical and biophysical research. The main applications of fluorescence can be found in DNA sequencing, imaging microscopy, and fluorescence spectroscopy.

The physical process of fluorescence can be best explained by a Jablonski diagram (shown in figure 18), named after Alexander Jablonski who also defined the concept of fluorescence anisotropy. The diagram shows the different energetic states of a single molecule and the transitions between those states.

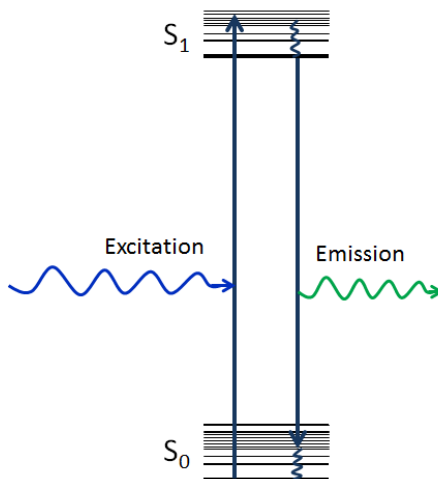


Figure 18: Jablonski diagram

Upon light illumination, a molecule can be excited from its ground state S_0 into one of its higher electronic states S_n , $n > 0$, by the absorption of a photon. All these electronic states contain vibrational energy levels. Dye molecules can be excited to higher vibrational energy levels by the absorption of photons in the visible spectrum. A possible return to the ground state is achieved by emission of a photon (fluorescence). The quantum yield ϕ of a molecule is the likelihood that a photon is emitted. It is defined as the number of radiative decays per absorbed photons (shown in equation 12).

$$\phi = \frac{n_{Fl}}{n_{Abs}} \quad (12)$$

The quantum yield can also be expressed in terms of rate constants (shown in equation 13).

$$\phi = \frac{\kappa_{Fl}}{\kappa_{Fl} + \kappa_{NR}} \quad (13)$$

In this formula, κ_{Fl} is the rate constant of the radiative transition (fluorescence) and κ_{NR} is the rate constant of non-radiative transition. The average time a molecule stays in the excited S_1 state until it returns to the S_0 state is called the fluorescence lifetime τ_{Fl} . Transition to the ground state can occur in different vibrational energy levels of S_0 . Fluorophores show typical lifetimes of several nanoseconds. The lifetime of a fluorophore can also be expressed in terms of rate constants according to equation (14):

$$\tau_{Fl} = \frac{1}{\kappa_{Fl} + \kappa_{NR}} \quad (14)$$

Therefore, the fluorescence lifetime τ_{Fl} depends on the quantum yield ϕ in the following way (shown in equation 15):

$$\tau_{Fl} = \frac{1}{\kappa_{Fl}} \cdot \phi \quad (15)$$

Emitted light usually has a longer wavelength and thus lower energy than absorbed light. This principle was first observed by Sir George Gabriel Stokes in 1852 and is called the Stokes shift. The main reason for the Stokes shift is the fast decay to the lowest vibrational level in S_1 . Transition to the ground state can occur into different vibrational energy levels of S_0 . The Stokes shift can be found in all kinds of fluorescent molecules. The fact that excitation and emission occur at different wavelengths is used to separate excitation from emission light by optical filters.

The absorption and emission spectra for most fluorophores are mirror images of each other, (shown in figure (19)).

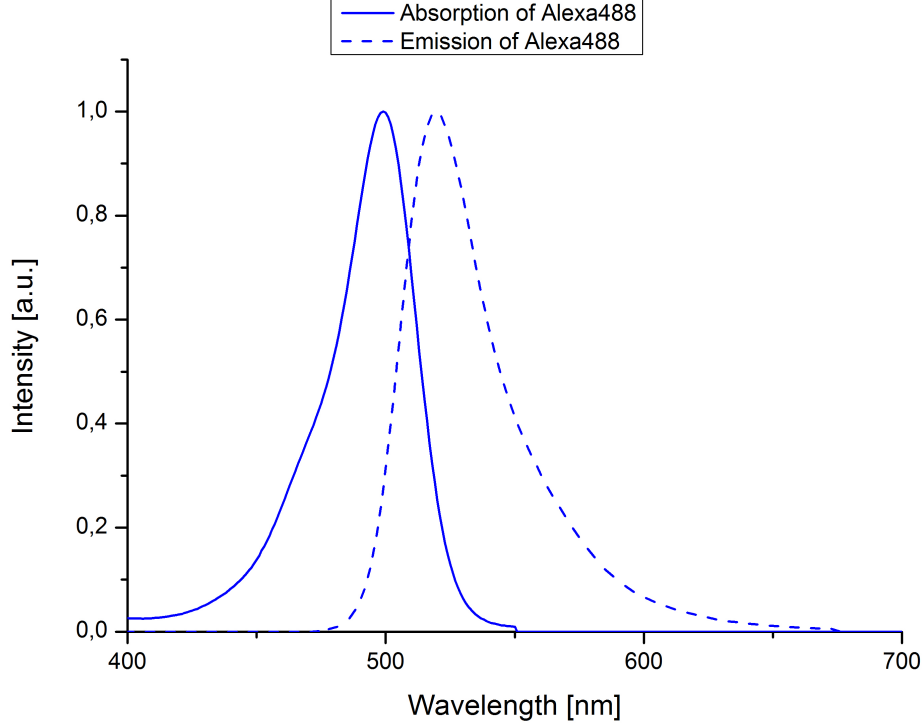


Figure 19: Stokes shift of Alexa Fluor 488. Absorption spectra of Alexa 488 in solid blue line, Emission spectra in dashed blue line.

In order to describe the spatial dependence of the fluorescence emission, light-emitting molecules can be seen as oscillating electric dipoles. To describe a dipole, it is necessary to know its charge q , the oscillation frequency ν , as well as the orientation and amplitude of the oscillation that are described by vector \mathbf{a} . The amplitude of the dipole moment \mathbf{p}_d is defined as the product of q times vector \mathbf{a} . According to the classic theory of Hertz, an electric field around the dipole can be written as follows (16):

$$\mathbf{E}_d(\mathbf{r}) = \left[\left(-\frac{k^2}{r} - \frac{3ik}{r^2} + \frac{3}{r^3} \right) \mathbf{e}_r (\mathbf{e}_r \cdot \mathbf{p}_d) + \left(\frac{k^2}{r} + \frac{ik}{r^2} - \frac{1}{r^3} \right) \mathbf{p}_d \right] \frac{e^{ikr}}{\epsilon} \quad (16)$$

In equation, (16) \mathbf{r} is the distance vector from the dipole while r is its modulus.

$$k = \frac{\omega}{c} = \frac{2\pi}{\lambda} \quad (17)$$

k is the modulus of the wave vector with ω being the circular frequency of the dipole's oscillation. $\omega = 2\pi\nu$, c is the speed of light in vacuum, \mathbf{e}_r is a unit vector pointing from the dipole towards the position r , and ϵ is the dielectric constant of the medium (shown in equation (18)). If the distance is significantly smaller than λ we find:

$$\frac{1}{r^3} > \frac{k}{r^2} > \frac{k^2}{r} \quad (18)$$

So in this range equation (eq:nearfield) can be approximated as:

$$\mathbf{E}_d(\mathbf{r}|r \ll \lambda) \approx [3\mathbf{e}_r(\mathbf{e}_r \cdot \mathbf{p}_d) - \mathbf{p}_d] \frac{e^{ikr}}{\eta^2 r^3} \quad (19)$$

In equation (19), the dielectric constant ϵ has been replaced by the square of the refractive index η .

2.13.2 Theory of FRET

The energy transfer between dipole antennas is known from the work of Helmholtz from 1886. This concept of energy transfer between dipoles was later transferred to the electromagnetic interaction between fluorescent molecules. The idea of energy transfer between molecules by dipole-dipole interactions was first described by Theodor Förster (Förster, 1948^[19]), this is why it is called the **F**örster resonance energy transfer (FRET).

Stryer and Haugland published the first experimental evidence of FRET from ruler samples with varying lengths between donor and acceptor in 1967 (Stryer *et al.*, 1967^[20]). In their experiments, they investigated the distance dependence of energy transfer on FRET-labeled poly-L-proline molecules. They used different peptide lengths between 12 and 46 Å.

Nowadays, FRET is a widespread method which is used in various scientific fields such as material science, life science, and biophysics. FRET describes the transfer of energy from a donor molecule to an acceptor molecule whenever they are in close proximity to each other. Donor and acceptor can also be viewed as dipole antennas. The efficiency of energy transfer between these antennas is dependent on their resonance frequency, which is analogous to the emission wavelength of the donor on the one hand and the excitation wavelength of the acceptor on the other hand.

The energy transfer efficiency E describes the probability that a photon from the donor is transferred to the acceptor. In equation (20), r is the distance between donor and acceptor, and R_0 is the Förster distance of donor and acceptor. Due to near-field interactions, the transfer efficiency depends on the 6th power of the distance.

$$E = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \quad (20)$$

The Förster distance R_0 is defined as the distance between donor and acceptor where the chance that energy from the excited donor molecule is transferred to an acceptor molecule is 50%. R_0 depends on a number of molecular properties.

$$R_0^6 = \frac{9000(\ln 10)\kappa^2 Q_D}{128\pi^5 N_A \eta^4} J \quad (21)$$

In equation (21), κ^2 is the orientation factor, Q_D is the quantum yield of the donor, N_A is Avogadro's number, η is the refractive index of the medium, and J is the spectral overlap of donor and acceptor. The latter is given by:

$$J = \int_0^\infty F_D(\lambda)\epsilon_A(\lambda)\lambda^4 d\lambda \quad (22)$$

J has the dimension $[M^{-1}cm^{-1}nm^{-4}]$, whereas $F_D(\lambda)$ is the normalized dimensionless fluorescence emission spectrum of the donor, and ϵ_A stands for the extinction coefficient of the acceptor $[1/M \cdot cm^2]$ (shown in figure 22).

R_0 depends on the relative orientation of the dipoles towards each other. The orientation factor κ can be calculated applying either equation (23) or equation (24).

$$\kappa = (\cos \theta_T - 3 \cos \theta_D \cos \theta_A) \quad (23)$$

$$\kappa = (\sin \theta_D \sin \theta_A \cos \phi - 2 \cos \theta_D \cos \theta_A) \quad (24)$$

The arrangement of all angles is shown in figure (20). Angle θ_T lies between the transition dipole of the donor emission and the transition dipole of acceptor absorption. The angles θ_D and θ_A lie between these dipoles and the vector joining donor and acceptor. ϕ is the dihedral angle between dipoles and connecting vector.

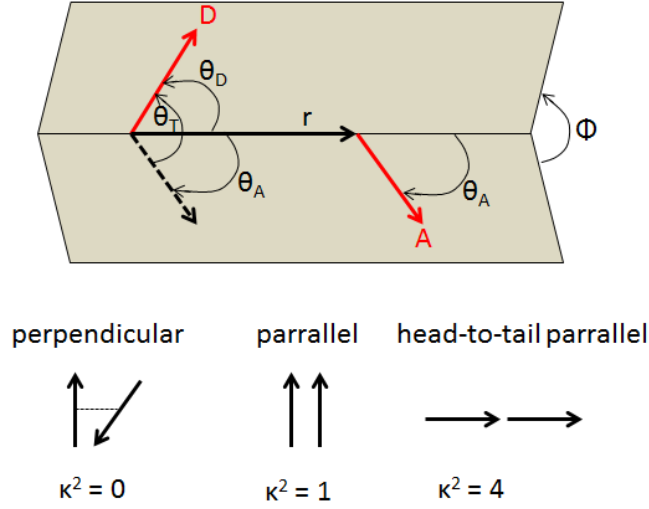


Figure 20: Orientation factor κ^2 . The transition dipole of donor (D) and acceptor (A) are shown in red.

The κ^2 values of the relative orientation range from 0 to 4. κ^2 is zero for any perpendicular orientation of dipoles, for parallel orientation κ^2 is 1 and for head-to-tail parallel orientation κ^2 is 4.

For a random rotational diffusion of fluorophores usually a κ^2 value of $2/3$ is used, which can be obtained by equation (25). The triangular brackets denote an averaging of all possible orientations.

$$\begin{aligned}
 \langle \kappa^2 \rangle &= \langle [2 \cos \theta_a \cos \theta_d - \cos \phi \sin \theta_a \sin \theta_d]^2 \rangle & (25) \\
 &= \langle 4 \cos^2 \theta_a \cos^2 \theta_d - 4 \cos \theta_a \cos \theta_d \cos \phi \sin \theta_a \sin \theta_d + \cos^2 \phi \sin^2 \theta_a \sin^2 \theta_d \rangle \\
 &= \frac{4}{9} + \frac{1}{2} \cdot \frac{4}{9} = \frac{2}{3}
 \end{aligned}$$

Anisotropy measurements can be performed to investigate the degree of free rotation of fluorophores.

By substituting the spectral properties of donor and acceptor as well as the donor quantum yield into equation (21), it is possible to calculate the Förster radii R_0 of FRET pairs.

The distance dependence of energy transfer is shown in figure (21). A strong distance dependence of energy transfer is found for r -values between $0.5R_0$ and $2R_0$. For $r = 2R_0$ the transfer efficiency has significantly decreased to 1.54%. Short distances between donor and acceptor, where $r = 0.5R_0$, lead to a transfer efficiency of 98.5%.

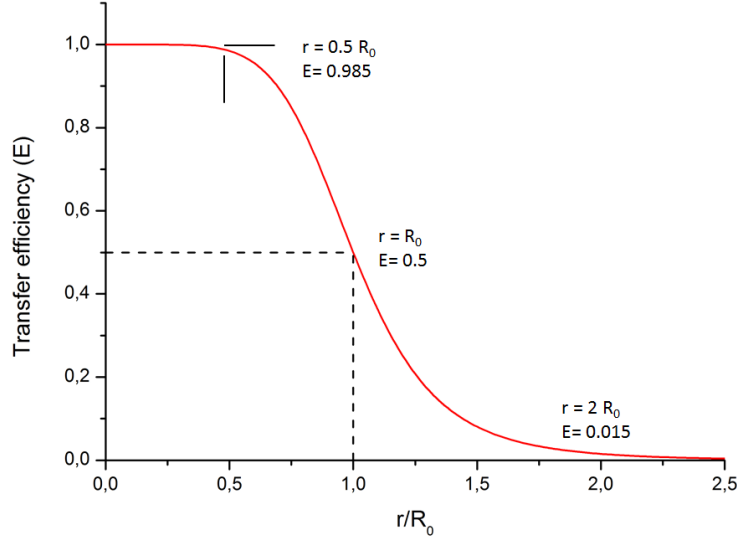


Figure 21: Transfer efficiency dependent on r/R_0

For typical donor and acceptor molecules, efficient energy transfer takes place on a distance range between 2 nm and 10 nm. Thus, FRET is an excellent tool to investigate cellular structures like lipid membranes which have a thickness of around 5 nm, or biomolecules like proteins and nucleic acids on a length scale of several nanometers. In most FRET studies, the main goal is to observe structural intra- and intermolecular changes of molecules. The distance r between two fluorophores can be calculated from measured FRET efficiency E , as shown in equation (20). With increasing FRET efficiency, the number of photons from the donor molecule decreases, which can be calculated using equation (26). The intensity-based analysis is more commonly used as compared to lifetime FRET analysis.

$$E = 1 - \frac{F_{DA}}{F_D} \quad (26)$$

There are two main methods for analyzing the transfer efficiency in FRET experiments: the first way is to measure the lifetime of the donor molecule. The lifetime of the donor shortens when energy is transferred to an acceptor. The second way to analyze transfer efficiency is to measure the intensities of donor and acceptor signal (shown in equation (27)).

$$\tau_{Fl} = \frac{1}{\kappa_{Fl} + \kappa_{NR} + \kappa_{FRET}} \quad (27)$$

Equation (28) shows how the quantum yield of the donor is reduced by the energy transfer to the acceptor molecule. Equations 26 and 27 show that energy transfer by FRET reduces the quantum yield ϕ , i.e. the fluorescence lifetime of the donor τ_{Fl} .

$$\phi_{Fl} = \frac{\kappa_{Fl}}{\kappa_{Fl} + \kappa_{NR} + \kappa_{FRET}} \quad (28)$$

Due to its distance and time resolution, FRET perfectly complements other techniques used for the analysis of protein structure and dynamics, e.g. NMR and x-ray crystallography.

In ensemble FRET measurements, many FRET pairs are measured simultaneously. The advantage of ensemble FRET measurements is a high number of photons collected per time due to the large number of excited molecules. FRET efficiencies measured by ensemble FRET show the average FRET efficiency of all measured molecules.

2.13.3 Selection of FRET pairs

The selection of donor and acceptor fluorophores to form a FRET pair is based on their R_0 value. It is known from X-ray crystallography that the diameter of the folded hPin1-WW-domain is around 24 Å. Therefore, fluorescent dyes were chosen whose R_0 values cover the entire size range of the folded and unfolded hPin1-WW-domain. Alexa Fluor 488 was selected as donor dye because its excitation maximum exactly fits the excitation wavelength of the blue lasers (470 nm) of the used setup. The other part of the FRET pair had to be an acceptor dye with an absorption maximum in the red to far red region so as to get a short R_0 distance caused by low spectral overlap. As a first option the FRET pair Alexa Fluor 488 and Alexa Fluor

750 was considered, which has a theoretical R_0 value of 41 Å. Another possibility was the FRET pair Alexa Fluor 488 and Alexa Fluor 647, which has a theoretical R_0 value of 56 Å. In theory, both FRET pairs can be used to measure folding-unfolding transitions of the hPin1-WW-domain. For single molecule measurements it is crucial to have bright molecules in order to get a sufficient signal-to-noise ratio. Therefore, the brightest of these two potential acceptor dyes needed to be used for labeling. The brightness was determined by a series of measurements in the fluorescence spectrometer (shown in figures 22 and 23).

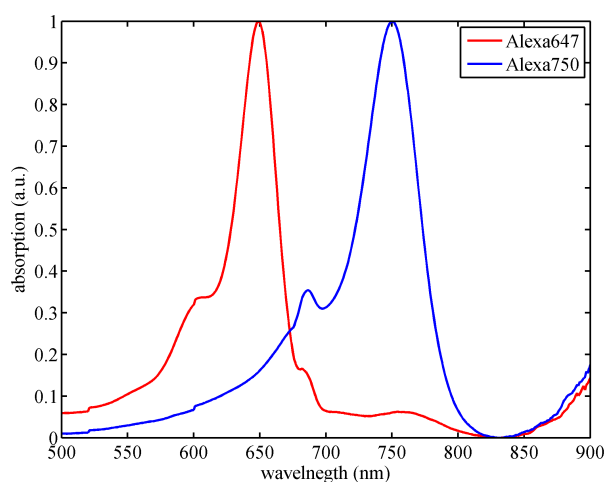


Figure 22: Absorption spectra of Alexa 647 (red) and Alexa 750 (blue)

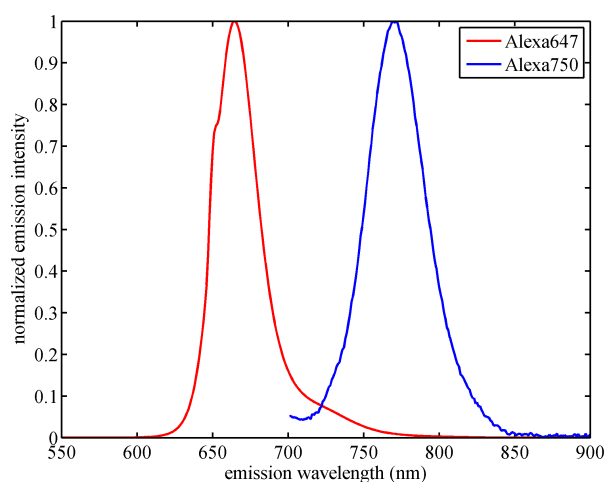


Figure 23: Normalized emission spectra of Alexa 647 (red) and Alexa 750 (blue)

The measurements indicated that the free Alexa Fluor 647 dye was roughly 22 times brighter than the free Alexa Fluor 750 dye. Thus, the Alexa Fluor 647 was selected as acceptor.

This dye has two negative net charges when it is attached to peptides. This difference in charge was used to purify and separate different species of labeled hPin1-WW-domain by ion exchange chromatography.

2.13.4 Acceptor labeling

The C- and N-terminal cysteine residues of the protein were labeled with fluorophores using maleimide chemistry. The mechanism of maleimide labeling works as the following: Dithiothreitol (DTT) breaks the disulfide bonds formed by cysteines. Thereby, generating free nucleophilic sulfhydryl groups. These nucleophilic groups are able to attack the double C-bond of the maleimide group during the labeling reaction, leading to the formation of a thioether group (schematic illustration shown in figure (24)).

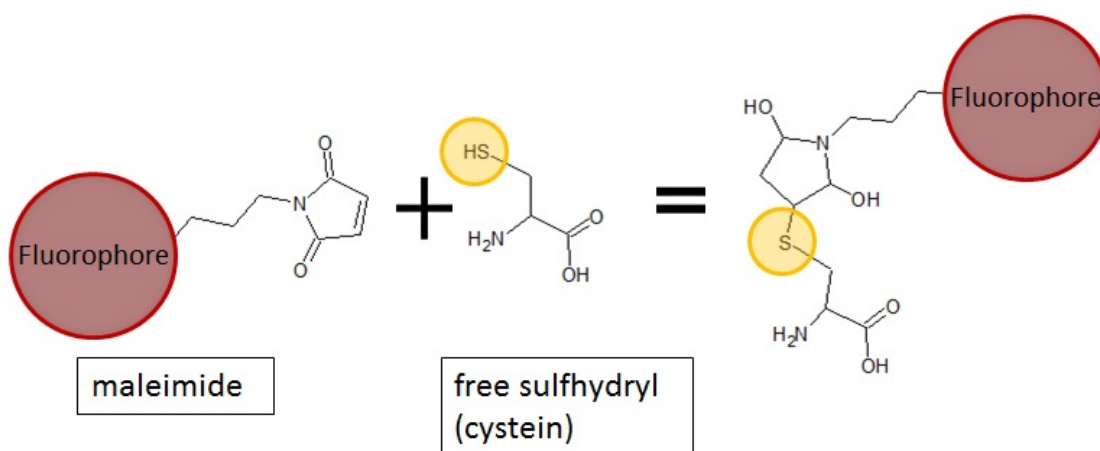


Figure 24: Schematic illustration of maleimide labeling reaction. The sulfur is emphasized in yellow.

The reduction was done using 10 mM Dithiothreitol (DTT) for 1 h at room temperature. Size exclusion chromatography was carried out to separate hPin1-WW-domain from GST, Thrombin and DTT. A sephadex column with a separation range of 100 Da to 7000 Da with PBS and a flow velocity of 0.4 ml/ min was used. The reduced cysteines of the hPin1-WW-domain were stochastically labeled using Alexa Fluor 647 maleimide. In order to make singly acceptor labeled molecules stochastically favorable, the labeling ratio of Alexa Fluor 647 to hPin1-WW-domain was

kept at 1 : 1. The labeling reaction was performed in PBS buffer at pH 7.4 for 6 h at room temperature. Afterwards, the concentration of the labeled protein was determined using a photometer (Jasco, Jena, Germany) at the maximum absorption of the Alexa Fluor 647 dye (650 nm).

2.13.5 Ion-exchange purification

Ion exchange chromatography was used to separate the proteins possessing one Alexa Fluor 647 dye from proteins with two or no dye molecule(s). Once attached to a protein, Alexa Fluor 647 has two negative charges. The difference in net charge of the proteins made it possible to separate the singly labeled hPin1-WW-domain from other fractions. The theoretical isoelectric point (pI) of the hPin1-WW-domain is 8.82. The buffer surrounding the protein was exchanged to 20 mM ammonium acetate at pH 5.0 using a PD10 column. At pH 5.0, the protein reached a positive net charge which was necessary to have the protein bind to the Hitrap SP HP Ion exchange column (GE, Frankfurt, Germany). Then, the column was equilibrated with eluent acetate buffer 20 mM pH 5.0. After the proteins had been bound to the matrix, a gradient of eluent acetate buffer 20 mM, 1 M NaCl pH 5.0 was used to elute the proteins from the column. The proteins bound to the matrix were competitively displaced by sodium cations. The first elution fraction was the doubly labeled hPin1-WW-domains which contained two dyes and therefore had the weakest affinity of all fractions. The next fraction corresponded to the singly labeled hPin1-WW-domain and the last fraction consisted of the unlabeled hPin1-WW-domain.

2.13.6 Donor labeling

The second step of labeling consisted of reducing the acceptor labeled hPin1-WW-domain by 10 mM Dithiothreitol (DTT) for 1 h at room temperature. Afterwards, the buffer of the reduced hPin1-WW-domain was exchanged for PBS pH 7.4 using a PD10 column, thereby removing free DTT. Again, the concentration of acceptor-labeled hPin1-WW-domain and donor dye Alexa Fluor 488 was determined using a photometer. The ratio of Alexa Fluor 488 to hPin1-WW-domain was kept at 4 : 1. A large excess of dye was necessary to allow for quantitative labeling of the protein. The labeling reaction was performed for 6 hours at room temperature. Finally, the double-labeled hPin1-WW-domain was separated from free Alexa Fluor 488 dye using a PD10 column.

2.13.7 Measurements in the fluorescence photometer

A fluorescence spectrophotometer (FP) (Jasco FP 6500, Jena, Germany) was used to measure FRET efficiencies of bulky protein solutions. Performing FP measurements, it is possible to excite molecules at a certain wavelength and measure their entire emission spectra. In this case, the concentrations of molecules used were in the nano- to micro-molar range. The concentration of hPin1-WW-domain was kept at 100 nM. Measurements were performed using a 3 mm quartz cuvette (Hellma). For every buffer condition, FRET efficiencies were measured by exciting the donor molecule Alexa Fluor 488 at 490 nm. Furthermore, emission of the donor was measured at 515 nm and the emission of the acceptor at 665 nm. As a control experiment, a donor-only sample of hPin1-WW-domain with Alexa Fluor 488 was measured under the same buffer conditions. When applying FP for FRET studies, it is possible to excite the donor molecule and measure the emission of the acceptor molecule to determine the FRET efficiency. Thus, it is possible to perform unfolding experiments with FRET samples. This unfolding can either be triggered by heat or by chemical denaturants like guanidinium chloride or urea. A concentration series was carried out to investigate the dependence of hPin1-WW-domain unfolding on guanidinium chloride. A range from 0 to 6 M guanidinium chloride was investigated. Thereby, it was possible to calculate the influence of guanidinium chloride on the fluorescence intensity of the dye itself.

2.13.8 Single molecule spectroscopy

Single molecule spectroscopy - in contrast to bulk measurements - investigates properties of single molecules. One advantage of this procedure is that it reveals the distribution e.g. of FRET efficiencies from heterogeneous ensembles of molecules. A second advantage is that artifacts arising from averaged signals can be avoided. One requirement for this method is a detection system which offers high sensitivity to light intensity and at the same time guarantees high time resolution. The optics of a confocal microscope reduce the out-of-focus light coming from the illuminated sample. Therefore, emission will be detected from a small detection volume called confocal volume. This results in an increase of contrast and resolution and a reduction of scattered light. Single molecule FRET experiments can be used to investigate protein folding dynamics and intra-/intermolecular interactions. Thus, it is possible to detect intramolecular distances which might be averaged out in ensemble FRET

measurements. Low concentrations of FRET pairs are used to avoid measuring more than one molecule at the same time. Therefore, the number of collected photons per time is usually lower than in ensemble measurements. The first single molecule FRET experiment was published by Ha and coworkers in 1996 (Ha *et al.*, 1996^[21]). **Pulsed interleaved excitation (PIE)** is used to excite molecules alternately with pulsed lasers. Thereby, it is possible to analyze lifetime and anisotropy of fluorescent molecules. PIE was first published by Lamb and coworkers in 2005 (Mueller *et al.*, 2005^[22]). As explained in their paper, fluorescent molecules are excited by interleaving excitation pulses. The lifetime of fluorophores is much shorter than the interval between excitation pulses. Thus, every emitted photon can be ascribed to a certain excitation pulse. The repetition period in PIE has a duration of 25 to 100 ns (shown in figure 25).

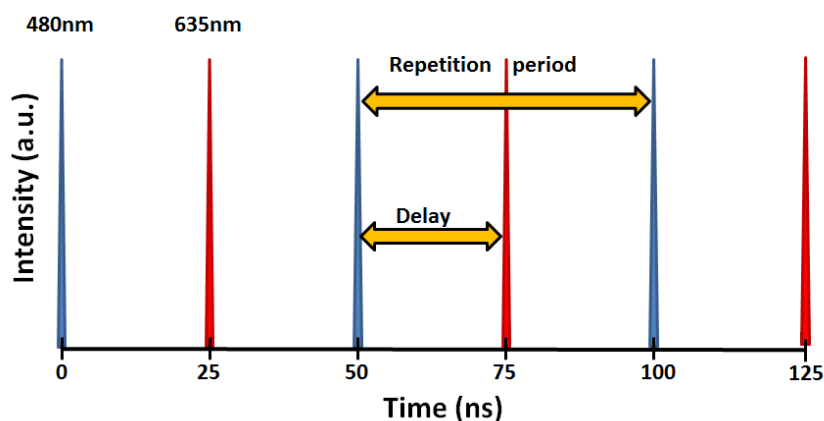


Figure 25: Pulsed interleaved excitation (PIE).

The determination of lifetimes can be valuable in FRET measurements. Therefore, **time correlated single photon counting (TCSPC)** is used which measures the arrival time of the emitted photons. The decay rate of relaxation can be calculated precisely by combining the arrival time of all photons emitted after the excitation pulse.

Finally, **multi-parameter fluorescence detection (MFD)** is a combination of measuring several fluorescence parameters, such as fluorescence lifetime (τ), fluorescence intensity (F), spectral wavelength (λ) and anisotropy (r). Thus, the information gained from single molecule measurements is maximized. MFD was first published by Seidel and coworkers (Eggeling *et al.*, 2001^[23]). The potential of MFD is the separation of molecule subpopulations in single molecule experiments.

A big challenge in smFRET experiments is the correction of measurement artifacts. Especially for a FRET pair with a high spectral overlap, direct acceptor excitation

and spectral crosstalk may arise. Direct acceptor excitation occurs from excitation of the acceptor by the (usually blue) laser which is used to excite the donor. In PIE measurements, this will lead to the emission of additional photons from the acceptor and these photons will then contribute to the measured FRET signal. The direct acceptor excitation α can be quantified by an acceptor-only measurement. α is defined as the ratio of the intensity from the blue laser I_D and the red laser I_A and the ratio of the absorption cross sections of the acceptor dye for the excitation by the blue laser $\epsilon_A^{\lambda_{exD}}$ and the red laser $\epsilon_A^{\lambda_{exA}}$ (shown in equation 29).

$$\alpha = \frac{\epsilon_A^{\lambda_{exD}} I_D}{\epsilon_A^{\lambda_{exA}} I_A} \quad (29)$$

The wavelength of direct acceptor excitation for Alexa Fluor 647 (shown in figure (26)).

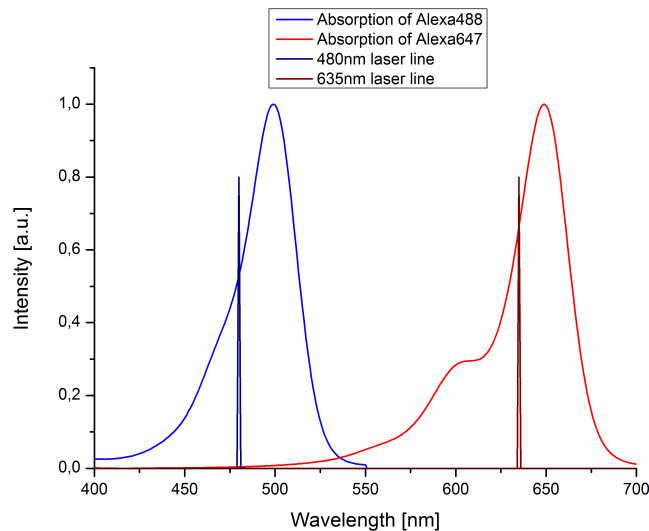


Figure 26: Direct acceptor excitation Shown are the absorption spectra of Alexa 488 (blue) and Alexa 647 (red) and the lasers used for excitation.

Spectral crosstalk refers to photons emitted by the donor that will be detected in the acceptor detection channel. In PIE measurements, these photons will increase the FRET signal because they are observed in the acceptor detection channel after the donor has been excited by the blue laser. The spectral crosstalk can be determined by a donor-only measurement and is quantified by the β -factor, which is defined in equation (30):

$$\beta = \frac{\eta_D^{\lambda_{emA}}}{\eta_D^{\lambda_{emD}}} \quad (30)$$

In equation (30) $\eta_D^{\lambda_{emA}}$ is the detection efficiency of the donor emission in the acceptor detection channel and $\eta_D^{\lambda_{emD}}$ is the detection efficiency of the donor emission in the donor detection channel.

Figure (27) shows that the emission of Alexa Fluor 488 in the red detection channel was very low which means that there was almost no spectral crosstalk in our FRET measurement.

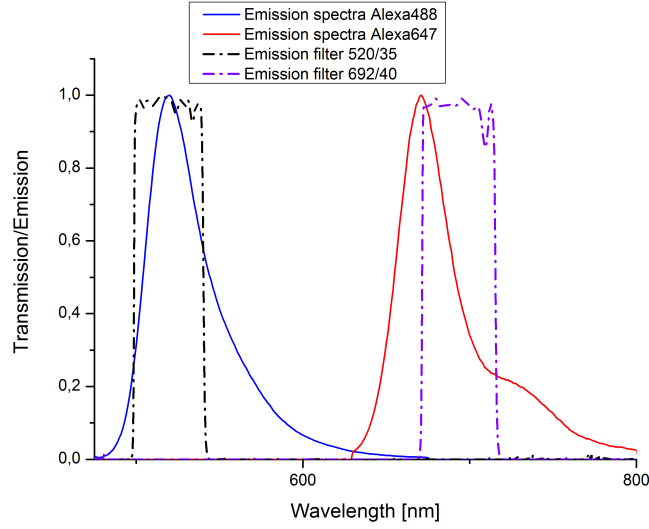


Figure 27: Spectral crosstalk. Shown are the emission spectra of Alexa 488 (blue) and Alexa 647 (red) and the properties of the emission filters (dashed lines).

A third important correction factor is the γ factor. It is used to correct the fluorescence quantum efficiency of donor ϕ_D and acceptor ϕ_A as well as the detection efficiency of donor $\eta_D^{\lambda_{emD}}$ and acceptor $\eta_A^{\lambda_{emA}}$ as defined in equation (31).

$$\gamma = \frac{\phi_A \eta_A^{\lambda_{emA}}}{\phi_D \eta_D^{\lambda_{emD}}} \quad (31)$$

Equation (32) shows the calculation of the corrected FRET efficiency using α -, β - and γ -factor. F_{BR} is the intensity detected in the acceptor channel after donor excitation, F_{RR} is the intensity detected in the acceptor channel after acceptor excitation and F_{BB} is the intensity detected in the donor channel after donor excitation.

$$E_{\text{corr.}} = \frac{F_{BR} - \alpha F_{RR} - \beta F_{BB}}{F_{BR} - \alpha F_{RR} - \beta F_{BB} + \gamma F_{BB}} \quad (32)$$

The stoichiometry determines the ratio of donor and acceptor molecules in the sample. Thereby, it is possible to identify and separate molecules without active donor or acceptor molecule from active FRET samples. A stoichiometry of 0.5 suggests an equal distribution of donor and acceptor molecules in the sample. The stoichiometry was calculated using equation (33).

$$S_{\text{corr.}} = \frac{F_{BR} - \alpha F_{RR} - \beta F_{BB} + \gamma F_{BB}}{F_{BR} - \alpha F_{RR} - \beta F_{BB} + \gamma F_{BB} + F_{RR}} \quad (33)$$

2.13.9 Setup

The setup used for all single molecule FRET measurement is shown in figure (28). It is based on a commercially available confocal microscope system (Microtime200, PicoQuant GmbH, Berlin, Germany). A standard inverted confocal microscope type IX71 (Olympus Deutschland GmbH, Hamburg, Germany) with a 60X water immersion objective 1.2 NA (UPLSAPO Olympus Deutschland GmbH, Hamburg, Germany) was utilized for the measurement. The excitation was carried out using two identical pulsed diode lasers, one for blue at 480 nm (LDH-P-C-485 PicoQuant GmbH, Berlin, Germany) and one for red at 635 nm (LDH-P-635 PicoQuant GmbH, Berlin, Germany). For the smFRET measurement, both lasers of the same color were synchronized. A single laser had a power of 100 μW , so that during the FRET measurement pulses were combined to 200 μW . The alternating laser pulses of blue and red laser light were generated by a PDL 828 diode laser driver (PicoQuant GmbH, Berlin, Germany). For smFRET measurements in PIE mode, the repetition rate was about 20 MHz with an interpulse distance of 50 ns (shown in figure 25). Regarding anisotropy measurements, two lasers of one color but perpendicular polarization were pulsed separately. A polarization conserving optical fiber was used to transmit the excitation light to a 10 x objective (UPLSAPO, Olympus) where it was collimated and then reflected by a dichroic mirror (51008 bs, Chroma Technology Corporation, Bellows Falls, USA) into the objective of the microscope. Emitted fluorescence light from the sample was collected by the same objective. A 150 μm pinhole was used as confocal aperture to decrease the background fluorescence.

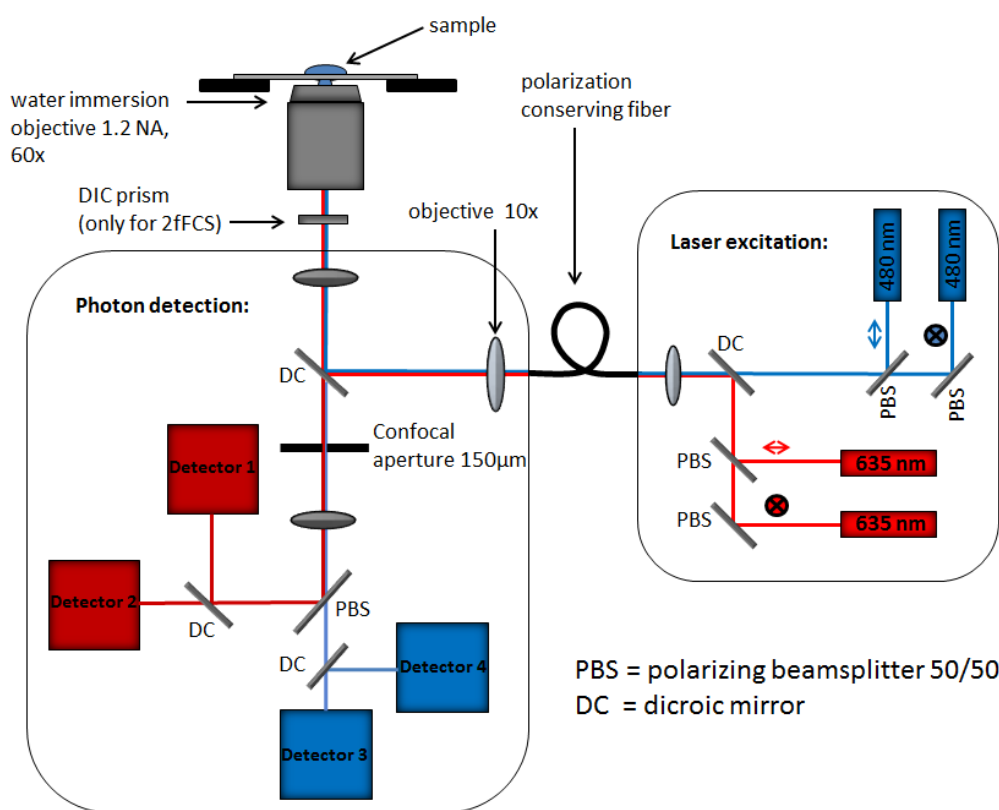


Figure 28: Microtime200 setup.

The passed light was collimated and split by a polarizing beam splitter and two dichroic mirrors (shown in figure 28). Each beam was then focused separately on single photon avalanche diodes (SPAD). Single molecule counting electronics (HydraHarp400, PicoQuant GmbH, Berlin, Germany) were used to record the arrival time of detected photons with a temporal resolution of 4 ps.

2.13.10 Single molecule FRET experiments

Measurements were taken in measurement chambers (Lab-tek, Nunc) which had a 170 μm cover slide at their bottom. The chambers were coated with Sigmacote (Sigma-Aldrich, Hamburg, Germany) for 45 minutes to avoid absorption of sample molecules to the walls. Four different measurements were performed for FRET analysis at each buffer condition (Kudryavtsev *et al.*, 2012^[24]). First of all, the pure buffer had to be measured for 20 min. Second, a 1 nM donor-only sample (single-labeled hPin1-WW-domain with Alexa 488) had to be measured for 20 min. Third, a 1 nM acceptor-only sample (single-labeled hPin1-WW-domain with Alexa 647) had to be measured for 20 min. At last, a 5 pM FRET sample (hPin1-WW-domain

with Alexa 488 and Alexa 647) had to be measured for 5 h to 10 h. The average number of molecules in focus was kept at a value around 0.1 to minimize the risk that more than one molecule would reside within the detection volume (Deniz *et al.*, 1999^[25]).

3 Results: Preparation of WW-domain FRET sample

This chapter describes the preparation of the WW-domain FRET sample. During the project, changes in strategy were necessary to reach the goal of a FRET-labeled molecule. Therefore, several variations of expression conditions, bacterial strains, plasmid vectors, and protein-variants were tested.

In a first step, the protein had to be expressed and then purified using an affinity-tag. Second, the affinity-tag had to be removed from the protein. Finally, the WW-domain had to be labeled with fluorophores.

3.1 WW-domain purification by His-tag

The first strategy involved the expression of four different types of WW-domains. All four WW-domains showed similarities in terms of 3D structure and conserved amino acids (red) in β -strand 1 and β -strand 3 (shown in figure 29):

	Strand 1	Strand 2	Strand 3
YAP65-WW-domain:	. . VPLPAGWEMAKTSS .	. GQRYFLNHID .	QTTT WQDPRKAM .
FBP28-WW-domain:	. . GATAVSEWTEYKTAD .	. GKTYYYNNRT .	LEST WE KPQELK .
YJQ8-WW-domain:	. . VRLPPGWEIHE . .	. NGRPLYNAEQ .	.KTKLHY P PSGS . .
WW-Prototype:	. . GLPPGWDEYKTH .	. NGKTYYYNHNT .	.KTSTWTDPRMSS .

Figure 29: Sequence alignment of WW-domains. Shown are the amino acid sequences of the Yes kinase associated protein 65 WW-domain from human (YAP65-WW-domain), the formin binding protein 28 from mouse (FBP28-WW-domain), the YJQ8-WW-domain from yeast and the prototype-WW-domain which is an artificial protein.

Several studies about different WW-domains have been published so far. Of those domains present in the literature four types were selected, thus forming a composition of WW-domains from different protein classes and various organisms. The idea was to analyze whether any distinction could be detected between the folding and unfolding of the four WW-domains, despite the fact that there were structural similarities between the proteins. The sequences of these proteins were taken from Macias (Macias *et al.*, 2000^[26]):

The WW-prototype is an artificial protein consisting of a mix of the sequence alignment of the three WW-domains.

DNA containing all four genes in one plasmid vector was produced by gene-synthesis (Invitrogen, Hamburg, Germany).

Each of the four genes had a specific restriction site surrounding it. By specifically cutting the genes from the plasmid vector, it was possible to separate them from each other. Table (13) shows which restriction endonuclease was used for the digestion of which gene:

Gene:	Restriction endonuclease:
YAP65-WW-domain	Xho I
FBP28-WW-domain	Bam HI
YJQ8-WW-domain	Sal I
WW-prototype	Xma I

Table 13: Restriction of WW-domain genes

The arrangement of the genes and their corresponding restriction sites in the plasmid vector (shown in figure (30)):

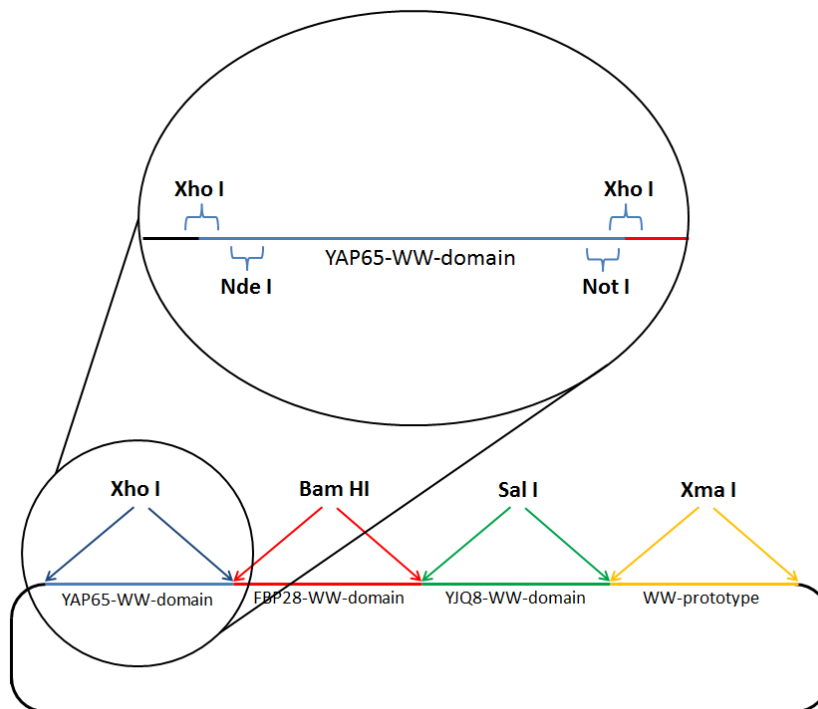


Figure 30: Restriction scheme of WW-domain genes for separation.

Subsequently, the genes of the four WW-domains were separated from the vector through applying agarose gel electrophoresis. Then, DNA bands of the isolated genes from YAP65-WW-domain, FBP28-WW-domain, YJQ8-WW-domain and the WW-prototype were extracted from the agarose. Visible were bands at ~ 5000 bp containing the plasmid vector pET24b (Novagen, Madison, USA), and bands at 120 bp containing the genes of the WW-domains. Shown in figure (31):

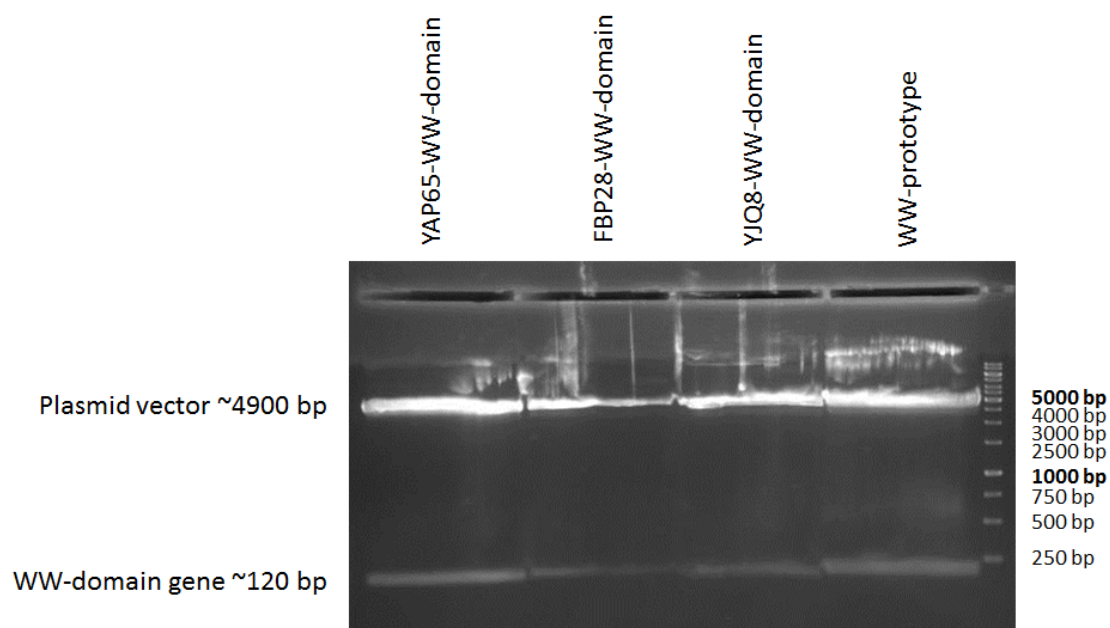


Figure 31: WW-domain genes after restriction in agarose gel stained by ethidium bromide.

Each of the four genes was framed by an Nde I and Not I restriction site in its DNA fragment, as shown in figure (30). Thus, it was possible to cut the four genes by Nde I and Not I restriction endonucleases from the plasmid vector. In the next step, the genes were ligated into the plasmid vector pET24b (shown in (56)), which had been cut by Nde I and Not I, too.

As a result of this ligation, the four genes were included in the plasmid vector pET24b (expression vectors 1), as shown in table (14).

YAP65-WW-domain-pET24b
FBP28-WW-domain-pET24b
YJQ8-WW-domain-pET24b
WW-prototype-pET24b

Table 14: Expression vectors 1

All four WW-domain proteins contained a histidine-tag (His-tag) for purification and a factor-Xa protease recognition site for removing the His-tag from the WW-domain after that purification. The schematic illustration of the WW-domains (shown in figure (32)).

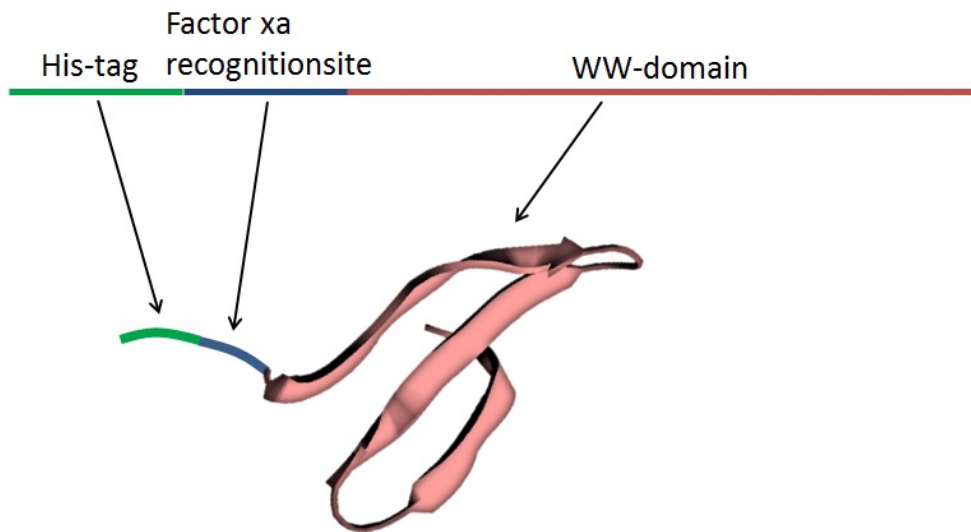


Figure 32: Scheme of WW-domain proteins.

For the expression of the proteins, each of the four expression vectors 1 was transferred separately into *E. coli* cells of the strain BL21 Gold by electrical transformation. Subsequently, the expression was done under the following conditions (time and temperature) shown in table (15):

20 °C	4 h - 32 h
25 °C	4 h - 24 h
30 °C	3 h - 24 h
37 °C	3 h - 16 h

Table 15: Expression conditions of WW-domains

The WW-domain proteins were purified by their His-Tag using a 5 ml Nickel-NTA (Ni-NTA) matrix (Invitrogen, Hamburg, Germany) in a self-packed column. PBS (pH 7.4) containing 300 mM imidazole was applied to elute the proteins.

An SDS-PAGE was carried out to monitor the expression efficiencies of the four WW-domain proteins. For this purpose, samples of each the YAP65-WW-domain, the FBP28-WW-domain, the YJQ8-WW-domain and the WW-prototype were analyzed in a 20% SDS gel.

There was no purified protein detectable in the SDS-PAGE. Possible reasons might have been that the protein was either not overexpressed or rather degraded in the cells, that it did not bind to the Ni-NTA matrix, or that it precipitated during purification. It was not possible to analyze what had happened to the protein.

The expression conditions had to be changed in order to achieve protein expression. Therefore, DNA of YAP65-WW-domain-pET24b, FBP28-WW-domain-pET24b, YJQ8-WW-domain-pET24b and the WW-prototype-pET24b was transformed into several *E. coli* strains which were believed to prevent degradation of overexpressed proteins in the cell (shown in table 16):

OverExpress(tm)C41(DE3)
OverExpress(tm)C41(DE3)plysS
OverExpress(tm)C43(DE3)
OverExpress(tm)C43(DE3)plysS
XL1-Blue

Table 16: Bacterial strains

Expression of the four expression vectors 1 was achieved applying the same variation of expression-time and expression-temperature as before (listed in table 15) in order to find optimum conditions. The expressed proteins were purified again by their

His-Tag in a Ni-NTA column. Samples were analyzed with SDS-PAGE. As in the previous attempt, no purified WW-domain could be detected in the SDS gel.

The genes of all four WW-domains were transferred into two other plasmid vectors in order to analyze the influence of the used plasmid vector on the expression.

All four genes were cut by the restriction endonucleases Nde I and Bam HI. They were ligated into the vectors pET11a (Novagen, Madison, USA 55) and pET27b (Novagen, Madison, USA 57) that had been cut by Nde I and Bam HI, too.

Thereby it was possible to obtain the following eight expression vectors 1 (shown in table 17):

YAP65-WW-domain-pET11a
FBP28-WW-domain-pET11a
YJQ8-WW-domain-pET11a
WW-prototype-pET11a
YAP65-WW-domain-pET27b
FBP28-WW-domain-pET27b
YJQ8-WW-domain-pET27b
WW-prototype-pET27b

Table 17: *Expression vectors 2*

Once more, the expression of all eight expression vectors 2 was achieved in every of the six bacterial strains used so far (16). Expression-time and expression-temperature were applied as before (listed in table 15) in order to find optimum conditions. Again, the expressed proteins were purified by their His-Tag in a Ni-NTA column and both expression and purification of the proteins were analyzed in a SDS-PAGE. Still, it was not possible to monitor purified WW-domain in the SDS gel.

A summary of bacterial strains, expression conditions, plasmid vectors and genes used is given in following table (18):

	BL21 Gold	OverExpress(tm)C41(DE3)	OverExpress(tm)C41(DE3)plysS	OverExpress(tm)C43(DE3)	OverExpress(tm)C43(DE3)plysS	XL1-Blue
20 °C	x	x	x	x	x	x
25 °C	x	x	x	x	x	x
Temperature 30 °C	x	x	x	x	x	x
37 °C	x	x	x	x	x	x

Table 18: Expression conditions of YAP65-WW-domain, FBP28-WW-domain, YJQ8-WW-domain and the WW-prototype in pET24b, pET11a and pET27b

It was impossible to express and purify a WW-domain protein under the conditions applied.

3.2 Calmodulin-WW-domain

The hypothesis was that the WW-domain might have been degraded inside the bacterial cell after the expression. Therefore, it was considered helpful to support the WW-domain expression and prevent it from degradation by fusing it to another protein. A fusion protein of calmodulin connected to the YAP65-WW-domain was produced in order to express and purify the WW-domain. For this purpose, the human Yes kinase associated protein 65 WW-domain (YAP65-WW-domain) was linked to calmodulin. The linker between calmodulin and WW-domain contained an enterokinase protease recognition-site where the protease was supposed to cut the WW-domain protein after purification. A schematic illustration of the fusion protein is shown in figure (33):

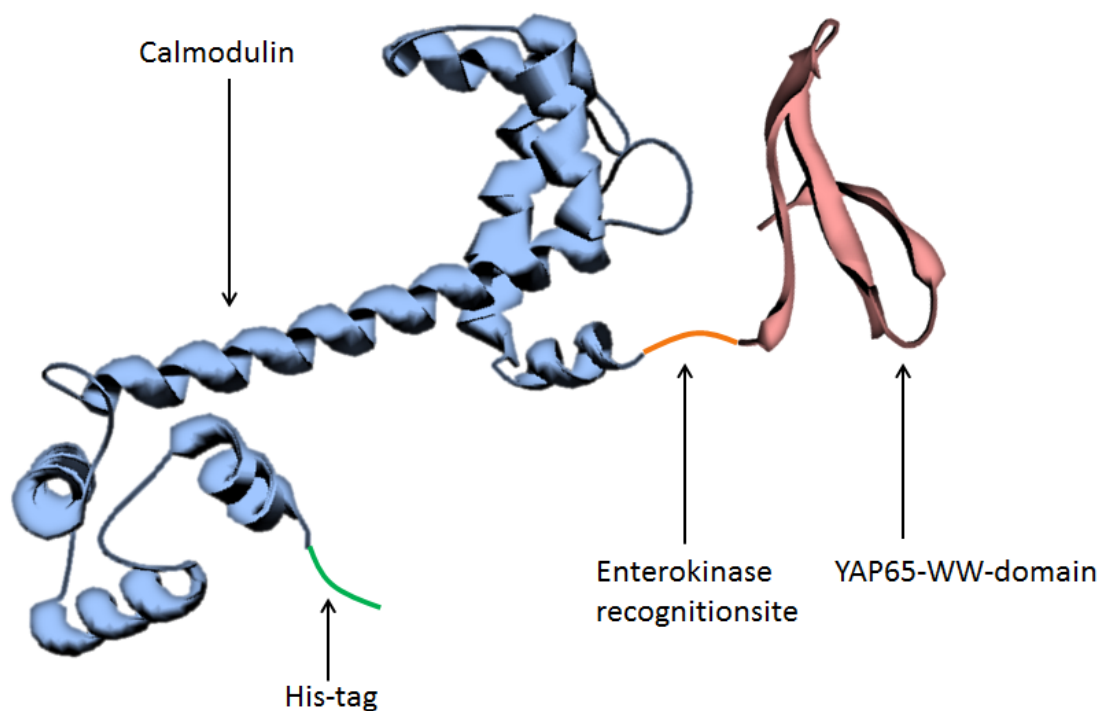


Figure 33: Schematic illustration of the calmodulin-YAP65-WW-domain.

In order to combine both proteins to one fusion protein, it was necessary to ligate the calmodulin gene and the YAP65-WW-domain gene together into the pET24b plasmid vector. The calmodulin gene had to be amplified and then modified by a PCR reaction in order to fuse it to the WW-domain. In order to do so, the calmodulin gene (which was available in our lab) was amplified with a linker region that connected it with the YAP65-WW-domain. The three-component ligation of the

calmodulin-WW-domain fusion protein is shown in figure (34). The PCR fragment of calmodulin was framed by a restriction site for the Nde I restriction endonuclease at the 5'-end and a restriction site for the Bam HI restriction endonuclease at the 3'-end in order to ligate it with the YAP65-WW-domain gene and the pET24b vector. Afterwards, the YAP65-WW-domain gene was cut with the restriction endonucleases Bam HI and Not I to ligate it to the calmodulin gene and the pET24b vector. Furthermore, the plasmid vector pET24b was cut with the restriction endonuclease Nde I and Not I in order to ligate it to the calmodulin gene and the YAP65-WW-domain gene. Subsequently, the three cut DNA fragments, i.e. calmodulin gene, YAP65-WW-domain gene, and pET24b plasmid vector, were ligated to one expression vector.

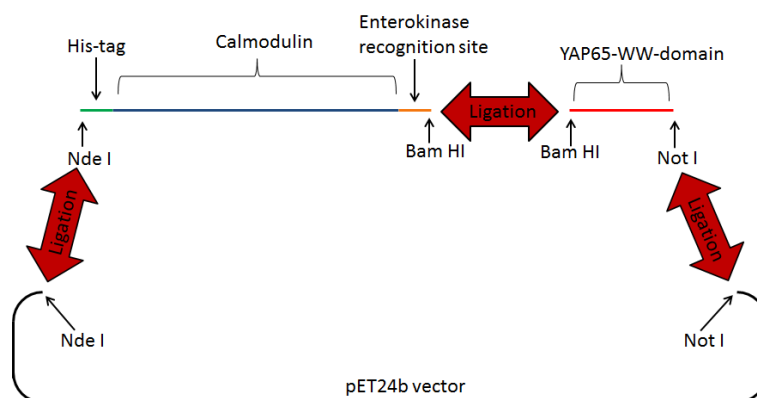


Figure 34: Ligation scheme of the calmodulin-WW-domain into the expression vector.

The resulting expression vector of the calmodulin-WW-domain fusion protein is shown in figure (35):

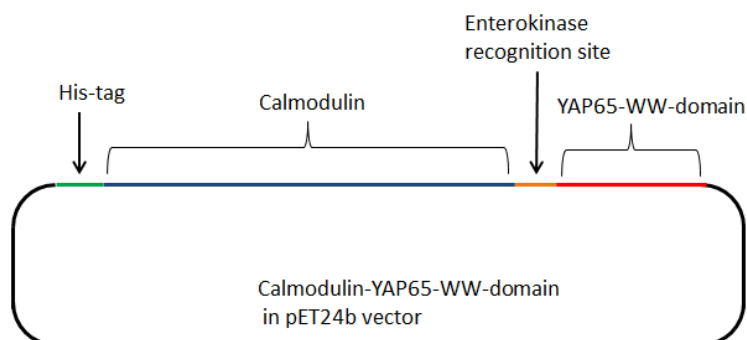


Figure 35: Expression vector of the calmodulin-WW-domain fusion protein.

The expression vector calmodulin-WW-domain-pET24b (CaM-WW-expression vector) was transformed into bacterial cells of the strain BL21 Gold.

Applying electrical transformation, the DNA of the calmodulin-WW-domain fusion protein (in the following referred to as: the fusion protein) was transferred into *E. coli* cells of the strain BL21 Gold. The expression conditions (time and temperature) were varied as shown in table (15).

The expressed fusion protein was purified by His-Tag using a Ni-NTA column matrix. The elution of the proteins was done with PBS (pH 7.4) containing 300 mM imidazole. Expression efficiency of the fusion protein was monitored via SDS PAGE in a 20 % SDS gel, which is shown in figure (36):

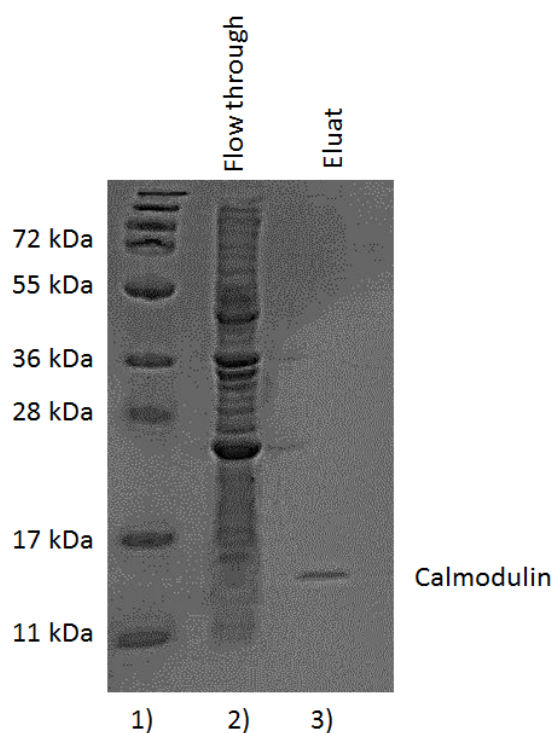


Figure 36: SDS gel of the Calmodulin-WW-domain fusion protein. The first line shows the protein ladder, the second line shows the flow through fraction and the third line shows the eluat which contains the calmodulin.

The gel contained two samples, namely the flow-through fraction of the calmodulin-WW-domain on the left lane 2 and the eluat fraction on the right lane 3. The flow-through fraction was supposed to contain non-specific bound proteins only, whereas the eluat fraction contained a protein band at 16 kDa corresponding to calmodulin without the WW-domain. It turned out that in being bound to calmodulin, the WW-domain had not been expressed.

Several other bacterial strains, which are listed in table (16), were used in order to find conditions where the entire fusion protein was expressed instead of only calmodulin.

Because it was still impossible to express and purify the fusion protein with the above mentioned methods, the experimental strategy had to be changed again.

3.3 GST-hPin1-WW-domain

The concept for the preparation of a hPin1-WW-domain FRET sample consisted in the application of a fusion protein to support the expression and purification of the WW-domain. A sequence alignment of the hPin1-WW-domain and the other WW-domains used in this work is shown in figure (37).

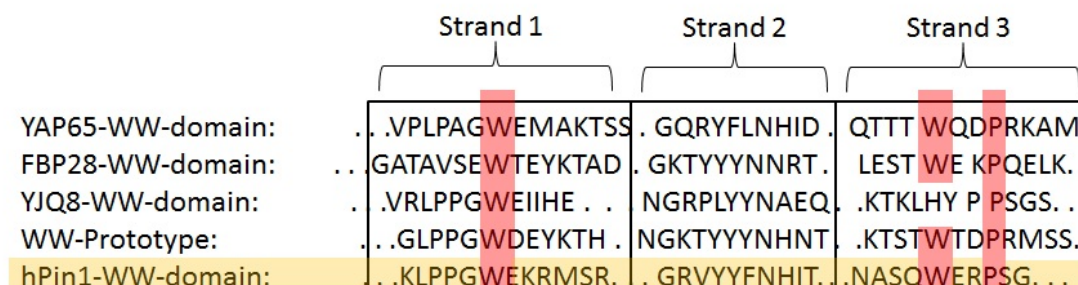


Figure 37: Sequence alignment of hPin1-WW-domain and other WW-domains. Shown are the amino acid sequences of the Yes kinase associated protein 65 WW-domain from human (YAP65-WW-domain), the formin binding protein 28 from mouse (FBP28-WW-domain), the YJQ8-WW-domain from yeast, the prototype-WW-domain which is an artificial protein and the rotamase WW-domain from human (hPin1-WW-domain)

In order to employ a fusion protein for the support of expression and purification, the used hPin1-WW-domain gene was ligated in a pGEX2T plasmid vector. Expression of a gene ligated into this vector led to a protein that was linked to a glutathione-S-transferase protein (GST) which was used as affinity tag to purify the fusion protein.

The hPin1-WW-domain used in this study was a W34F mutant. Like Liu *et al.* (Liu *et al.*, 2007^[27]) showed, there was no influence on structure and stability of the WW-domain caused by this mutation. Still, the gene of the hPin1-WW-domain had to be modified in order to allow FRET labeling. Therefore, two cysteines had to be added to the hPin1-WW-domain. A modification of the DNA was achieved by two QC-PCR reactions.

In a first step, a base triplet TGC that codes for a cysteine was added at the 5' region of the hPin1-WW-domain. The following primers QC-hPin1-WW C3 fwd and QC-hPin1-WW C3 rev were used for the QC-PCR amplification reaction (shown in figure 19):

QC-hPin1-WW C3 fwd:
AAT TCC CGG GGA TCC TGC GAA GAA AAG CTG CCG CCC GGC
QC-hPin1-WW C3 rev:
GCC GGG CGG CAG CTT TTC TTC GCA GGA TCC CCG GGA ATT

Table 19: QC-hPin1-WW C3 primers

The second step consisted of a QC-PCR inserting a TGC triplet coding for a cysteine at the 3' region of the hPin1-WW-domain gene (shown in figure 20).

QC-hPin1-WW C40 fwd:
CAG TTC GAG CGG CCC AGC GGC TGC CAG CTC CCG GAG ACG
QC-hPin1-WW C40 rev:
CGT CTC CGG GAGCTG GCA GCC GCT GGG CCG CTC GAA CTG

Table 20: QC-hPin1-WW C40 primers

For the second QC-PCR amplification reaction, the primers QC-hPin1-WW C40 fwd and QC-hPin1-WW C40 rev were used to mutate a TGC triplet into the hPin1-WW-domain-C3-pGEX2T DNA, which was used as template.

The QC primers (blue) bound to the template DNA of hPin1-WW-domain (black) in pGEX2T vector (red), and inserted a TGC codon at the 5'- and 3'-parts of the gene. In figure (38), the capital letters represent the amino acid sequence of the hPin1-WW-domain (black) after translation.

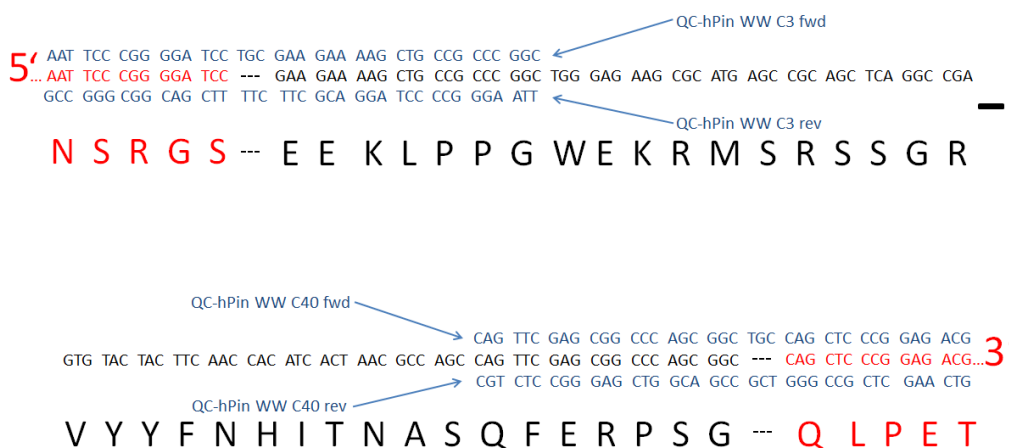


Figure 38: hPin1-WW-domain in pGEX2T vector with QC primers.

The DNA (black) and the amino acid sequence (capital letters, black) of the hPin1-WW-domain after insertion of TGC codons into the gene are shown in figure (39):

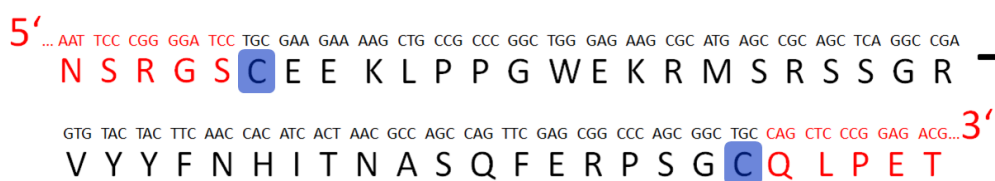


Figure 39: hPin1-WW-domain in pGEX2T with N- and C-terminal cysteines.

The mutated expression vector hPin1-WW-domain-C3C40-pGEX2T (in the following referred to as: the hPin1-WW-expression vector) was transformed into bacterial cells of the strain BL21 Gold.

Expression of the hPin1-WW-domain was conducted using BL21 Gold cells containing the hPin1-WW-expression vector.

3.3.1 SDS-PAGE

The progress of the proteolytic digestion of the GST-fusion protein by thrombin was monitored using SDS PAGE. For this purpose, 5 μl loading dye and 5 μl H_2O were added to the collected sample of the proteolytic digestions and loaded into a pocket of a 20 % SDS gel, as shown in figure (40):

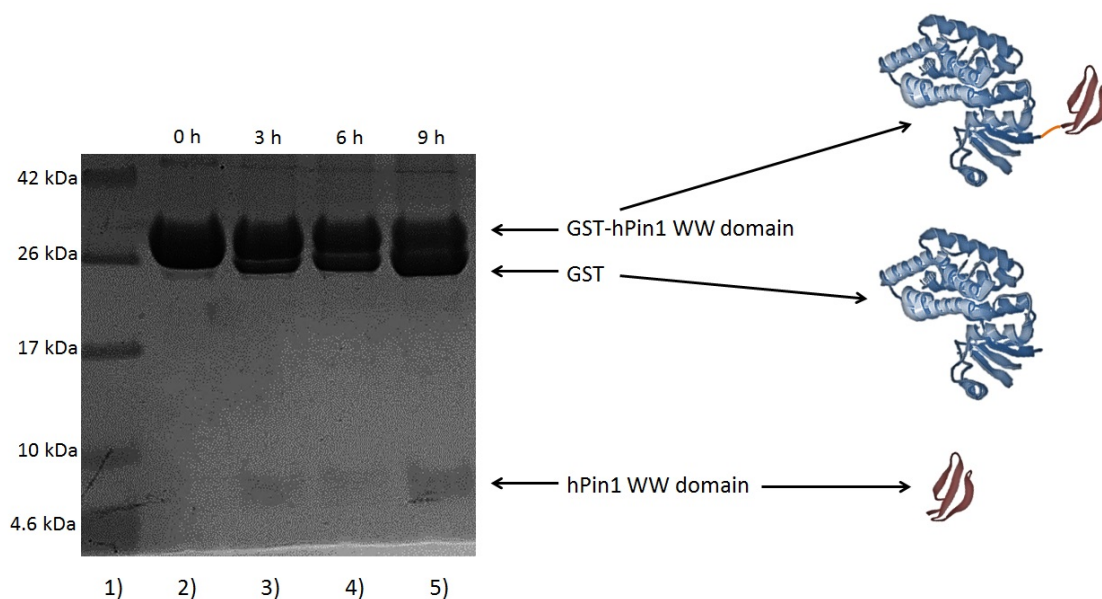


Figure 40: SDS-PAGE of GST-hPin1-WW-domain cut by thrombin. The first line shows the protein ladder, the lines two to five contains the GST-hPin1-WW-domain cut by thrombin with incubation times between 0 h and 9 h.

The SDS gel showed protein bands that corresponded to the calculated size of the GST-fusion protein (32-kDa), as well as protein bands that were the equivalent to the GST with a size of 26 kDa. Furthermore, it showed protein bands that corresponded to the hPin1-WW-domain with a size of 6 kDa. The increase in concentration of GST and hPin1-WW-domain over time corresponded to the progress of cleavage by thrombin. Therefore, the protein bands with the strongest signal of free GST and hPin1-WW-domain were found after 9 h incubation time. A precipitation of proteins was observed for those cleavage preparations that exceeded nine hours. Thus, the incubation time could not be longer than 9 h, although there was still uncleaved fusion-protein at that time.

3.3.2 HPLC purification

The thrombin-cut of the GST-fusion protein was incubated with 10 mM DTT at room temperature for 1 h to reduce the C- and N-terminal cysteine residues in order to prepare them for the labeling with fluorophores by maleimide chemistry. By reducing the protein before performing the HPLC purification, it was possible to separate the hPin1-WW-domain from other proteins and to remove the DTT in one step. After the DTT incubation, 1 ml of the GST-fusion protein solution from the proteolytic digestion with thrombin was used to perform a separation by HPLC. This purification was done using a Superdex Peptide column. During the HPLC run, a peak at 280 nm in the absorption spectrum indicated the proteins in the eluate. Analysis of the HPLC elution diagram data showed a very strong peak which was supposed to correspond to the uncut GST-fusion protein (32 kDa), GST (26 kDa), and thrombin (36 kDa) because these three proteins had a size above the column's void volume size of 20 kDa. Therefore, they were most likely contained in the cut-off fraction which was eluted first. The free hPin1-WW-domain was eluted in one of the next fractions. Five different molecule fractions were collected separately from each other, due to the separation conducted by the column (shown in figure 41).

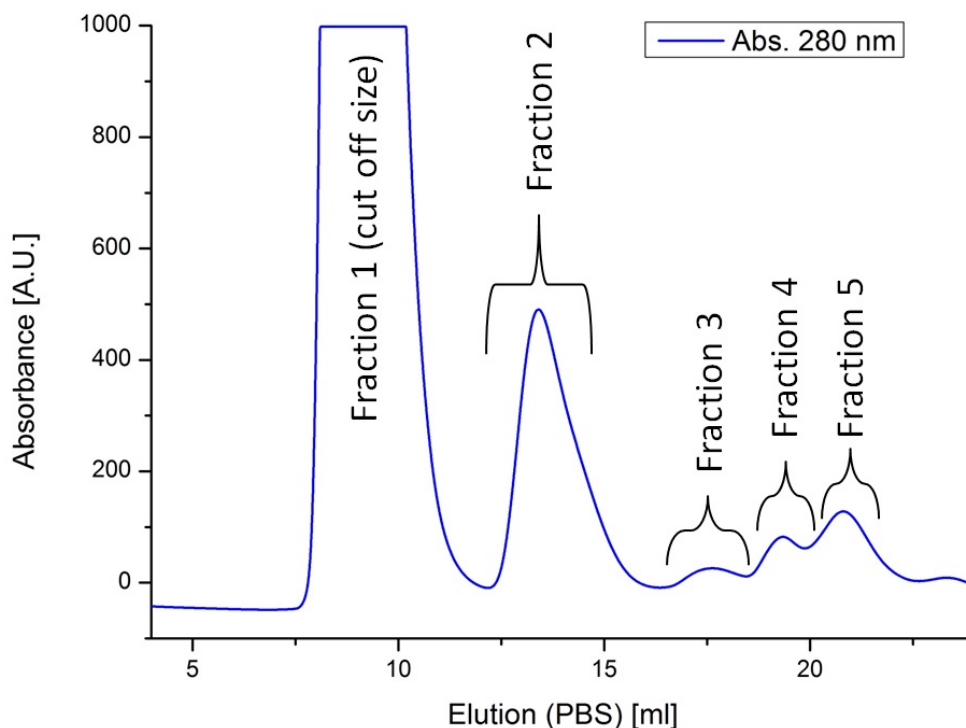


Figure 41: *hPin1-WW-domain purification*

Fraction 2 was collected from 12.5 ml to 15 ml and fraction 3 was collected between 16.5 ml and 18.5 ml. The hPin1-WW-domain had to be identified in one of the collected fractions. To do so, the tryptophan emission of the hPin1-WW-domain was measured according to chapter 2.13.7. Buried tryptophan residues have a higher emission in respect to an exposed tryptophan due to quenching effects ([28]). From the crystal structure, the hPin1-WW-domain was known to contain a tryptophan residue which is cached from the surrounding solvent.

The second and third of the collected protein fractions were measured under native conditions in PBS (blue) and unfolded in 6 M GdCl (red) in order to determine the location of the tryptophan.

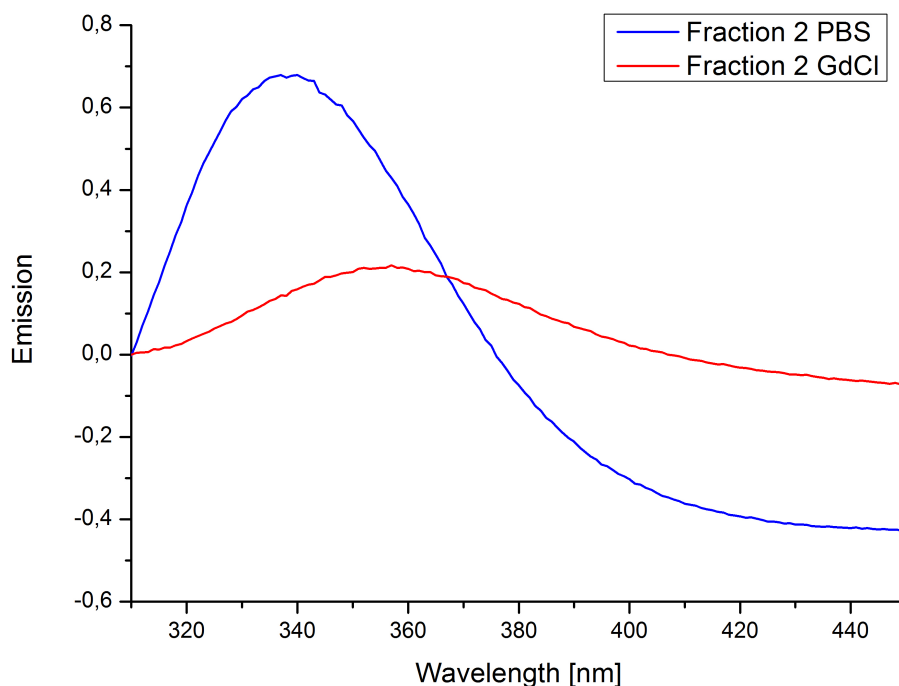


Figure 42: Identification of Fraction 2

In the unfolded state, the tryptophan's emission intensity from the protein of fraction 2 was much lower than in the folded state. This indicated that the protein in fraction 2 contained a buried tryptophan residue (shown in figure 42). Thus, it was concluded that fraction 2 contained the hPin1-WW-domain. The spectrum of the unfolded fraction 2 was shifted to the red spectral region, which is known to happen for tryptophan emission under high salt buffer conditions ([28]).

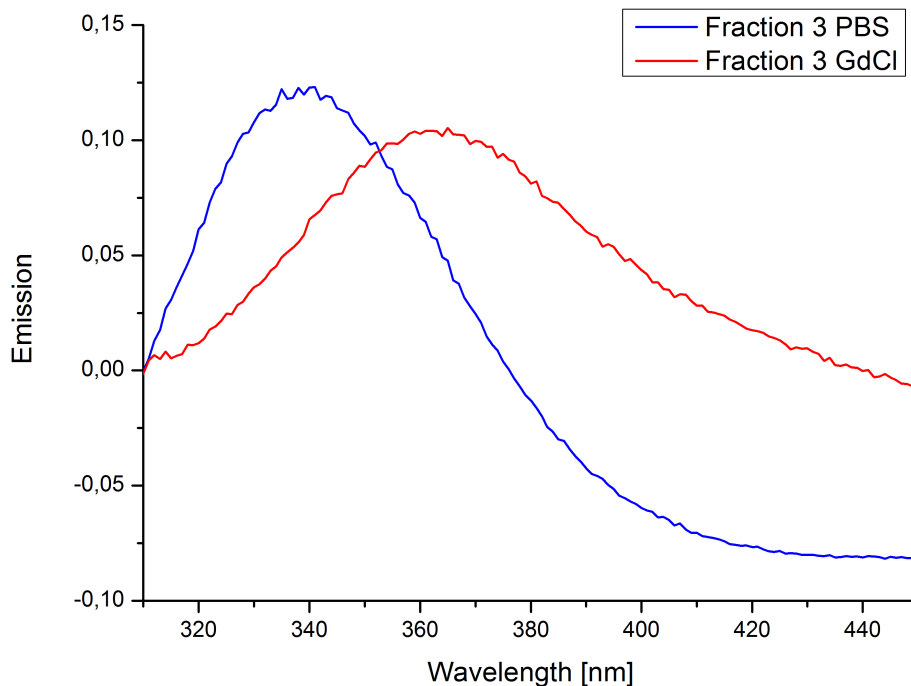


Figure 43: Identification of Fraction 3

On the other hand, the tryptophan emission intensity of fraction 3 showed no big difference between the folded and unfolded state (shown in figure 43). Here, the analysis did not show the typical emission behavior of a protein with a buried tryptophan.

3.3.3 Determination of protein concentration

The concentration of proteins was determined using the Lambert-Beer law (shown in equation 34):

$$c = \frac{A}{\epsilon \cdot l} \quad (34)$$

In this equation, c is the concentration of the sample [M], A is the absorbance, ϵ is the molar extinction coefficient of the sample [$M^{-1}cm^{-1}$], and l is the length of the cuvette [cm]. A photometer was employed to measure the absorbance of the hPin1-WW-domain.

The baseline for correction was measured at 320 nm and subtracted from the tryptophan absorbance of the hPin1-WW-domain which, in turn, had been measured at 280 nm (shown in figure 44). The calculated molar extinction coefficient of hPin1-WW-domain with reduced cysteine residues was $8730 \text{ M}^{-1}\text{cm}^{-1}$ at 280 nm.

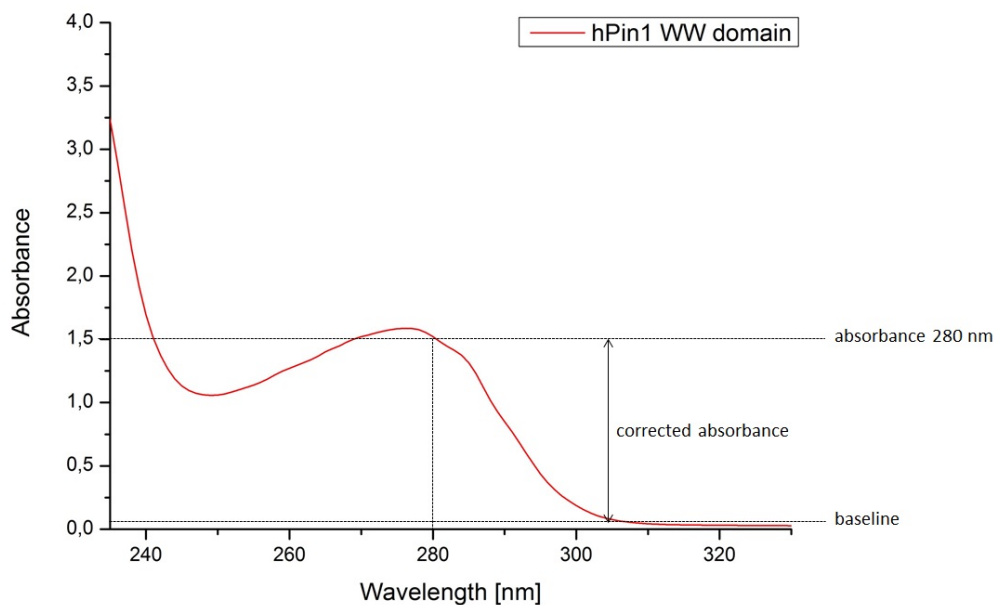


Figure 44: Determination of concentration of hPin1-WW-domain.

The protein concentration was calculated from the baseline-corrected absorbance at 280 nm divided by the molar extinction coefficient of the hPin1-WW-domain (shown in equation 35):

$$\text{Conc. protein} = \frac{A}{\epsilon \cdot d} = \frac{1.49}{8730 \text{ M}^{-1}\text{cm}^{-1} \cdot 1 \text{ cm}} = 171 \mu\text{M} \quad (35)$$

Consequently, the resulting solution had a concentration $171 \mu\text{M}$ and a total volume of 8 ml.

3.3.4 Mass spectrometry of proteins

The band of the GST-hPin1-WW-domain and of the hPin1-WW-domain were extracted from the SDS gel and analyzed in the lab of Dr Oliver Valerius in the department of microbiology at the University of Göttingen.

An evaluation of the analysis is shown in table (21).

File	Peptide	XC
hPin1	R.VYYFNHITNASQFER.P	3.5
hPin1	R.VYYFNHITNASQFERPSGC	3.07
GSThPin1	K.ERAEISMLEGAVLDIR.Y	3.38
GSThPin1	K.KFELGLEFPNLPYYIDGDVK.L	4.06
GSThPin1	K.YIAWPLQGWQATFGGGDHPPK.S	2.54

Table 21: Mass spectrometry evaluation of hPin1-WW-domain and GST-hPin1-WW-domain. The XC value is determined by the number of N-terminal and C-terminal protein fragments that were dedicated to each other, thereby identifying a certain protein fragment. A XC value above 2 indicates a significant number of dedicated N-terminal and C-terminal fragments.

The evaluation of the mass spectrometry analysis identified the hPin1-WW-domain and the GST-hPin1-WW-domain and thereby showed that both proteins were expressed and purified.

3.3.5 Acceptor labeling

The reduced cysteine residues of the hPin1-WW-domain were used for labeling with Alexa 647 maleimide. Maleimide groups are able to form a bond to thiol groups of cysteines. 100 μg of lyophilized Alexa Fluor 647 maleimide dye were dissolved in 50 μl PBS (pH 7.4). To measure its concentration, this solution was diluted 100 times with PBS (pH 7.4) and its absorbance spectrum was measured between 400 nm and 600 nm in the photometer. Furthermore, the concentration of the Alexa Fluor 647 was calculated using its molar extinction coefficient $270000 \text{ M}^{-1}\text{cm}^{-1}$ (Invitrogen, Hamburg, Germany) at its absorption maximum of 668 nm applying the Lambert-Beer law.

The absorbance multiplied by a factor of 100 due to the dilution was 453.6. The concentration of Alexa 647 was calculated using formula (36).

$$c_{\text{dye}} = \frac{A_{\text{Alexa647}}}{\epsilon \cdot l} = \frac{454}{270000 \text{ M}^{-1}\text{cm}^{-1} \cdot 1\text{cm}} = 1.7 \text{ mM} \quad (36)$$

The concentration of the hPin1-WW-domain was 171 μM . Labeling of the protein was achieved in the ratio of 1:1.

$$\text{Labeling conc.} = \frac{C_{\text{Alexa647}}}{C_{\text{hPin1-WW-domain}}} = \frac{1680\mu\text{M}}{171\mu\text{M}} \approx 10 \quad (37)$$

It turned out that the concentration of the Alexa Fluor 647 was 10 times higher than the concentration of the hPin1-WW-domain. Therefore, the 10 fold volume of the hPin1-WW-domain was added to 50 μl Alexa Fluor 647 in order to reach the labeling ratio of 1:1. This 1:1-ratio is stochastically favorable because thus only one of the two cysteine residues per protein is labeled (shown in equation 38).

$$\text{Labeling volume of hPin1-WW-domain} = 10 \cdot 50 \mu\text{l} = 500 \mu\text{l}. \quad (38)$$

Afterwards, the labeling mixture was incubated at room temperature for 3 hours.

3.3.6 Ion-exchange purification

Purification of the singly labeled hPin1-WW-domain Alexa 647 was achieved by applying ion-exchange chromatography. The elution diagram showed two peaks that were detected at 650 nm, as is shown in figure (45).

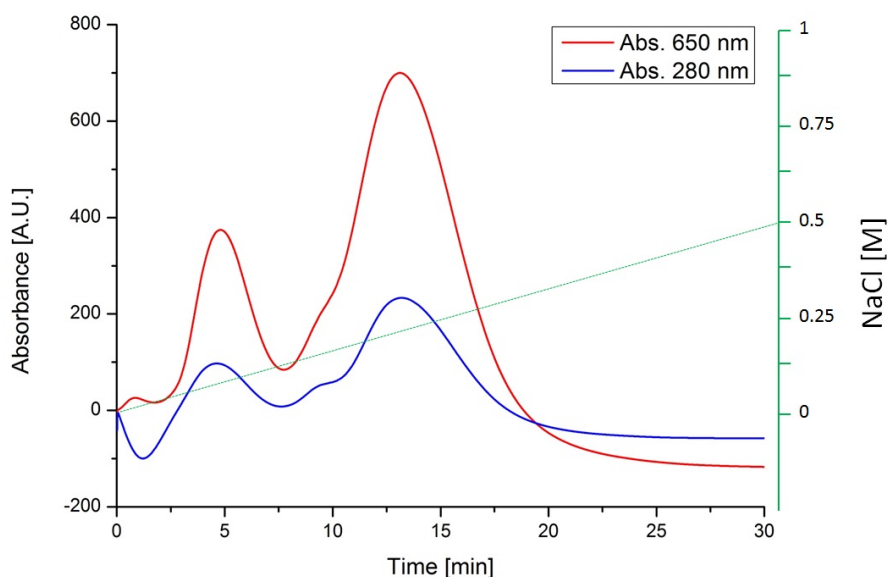


Figure 45: Ion exchange chromatography of hPin1-WW-domain Alexa Fluor 647

While the first peak corresponded to the hPin1-WW-domain doubly labeled with Alexa Fluor 647 dye, the second peak equated to the hPin1-WW-domain singly labeled Alexa Fluor 647 dye. A full absorption spectrum of the fraction collected between 12.5 ml and 15 ml was measured between 230 nm and 800 nm (shown in figure 46):

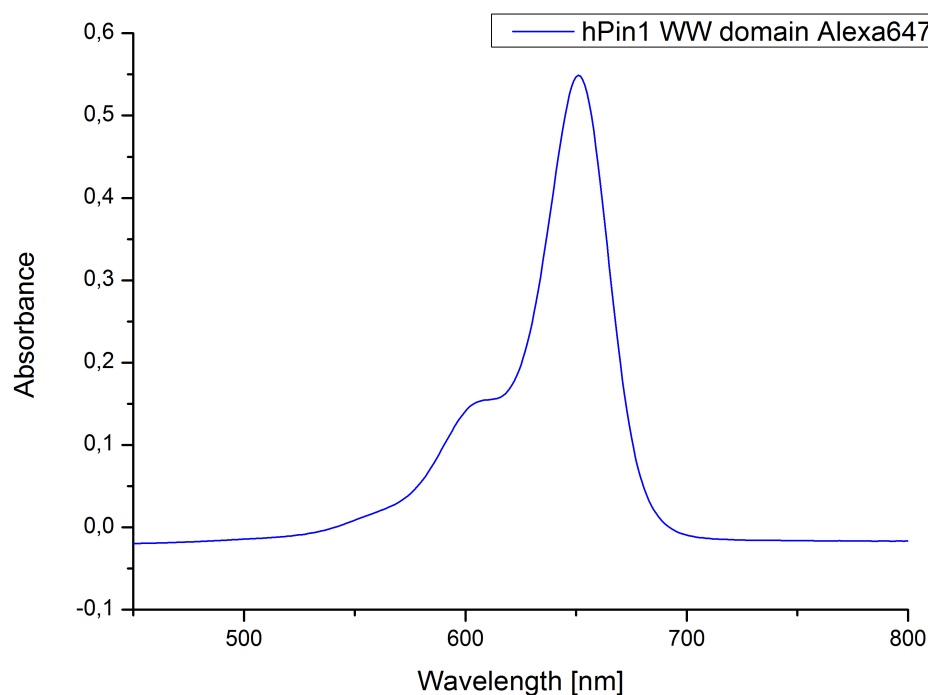


Figure 46: hPin1-WW-domain Alexa Fluor 647 spectrum

The concentration of the Alexa Fluor 647 absorption spectrum was determined at its absorption maxima of 668 nm. Also, the baseline value for the correction was measured at 750 nm. Applying equation (36), the concentration of the hPin1-WW-domain Alexa Fluor 647 was calculated to be 2.1 μM . The volume of the sample was 1.8 ml.

3.3.7 Donor labeling

In a next step, the hPin1-WW-domain labeled with a single Alexa Fluor 647 was labeled with the donor dye Alexa Fluor 488 in order to prepare a doubly labeled FRET sample. Having been dissolved in 50 μl PBS pH 7.4, the maximum absorption of the Alexa Fluor 488 dye was measured at 500 nm to determine its concentration. Equation (39) was used for this calculation. The baseline value for the correction was measured at 600 nm.

$$\text{Concentration Alexa Fluor 488} = \frac{22.8}{73000 \text{ M}^{-1}\text{cm}^{-1}} = 312 \mu\text{M} \quad (39)$$

The labeling of the protein was achieved with a four to five fold excess of Alexa Fluor 488.

It was found that the concentration of Alexa Fluor 488 was 150 times higher than the concentration of the singly labeled hPin1-WW-domain Alexa Fluor 647, while the total volume of the protein was 36 times higher than the total volume of the dye.

Fifty μl Alexa Fluor 488 dye were mixed with 1800 μl of the protein, thereby getting an excess of Alexa Fluor 488 of 4.1 : 1 with respect to the protein (shown in equation 40).

$$\text{Labeling ratio Alexa 488 : hPin1-WW-domain-Alexa 647} = \frac{150}{36} = 4.1 \quad (40)$$

The labeling mixture was incubated at room temperature for three hours according to chapter 2.13.4. Free dye was removed using a PD10 column, thereby exchanging the buffer of the double labeled hPin1-WW-domain to PBS pH 7.4. A determination of dye concentrations showed that there still was free donor dye left in the sample. To remove that free dye, the purification in the PD10 column was repeated.

The concentration of the double labeled hPin1-WW-domain was determined by measuring an absorption spectrum between 400 nm - 750 nm. For the correction, the baseline value was measured at 750 nm (shown in figure 47). Then, the concentrations of Alexa Fluor 488 and Alexa Fluor 647 were calculated according to equation (36).

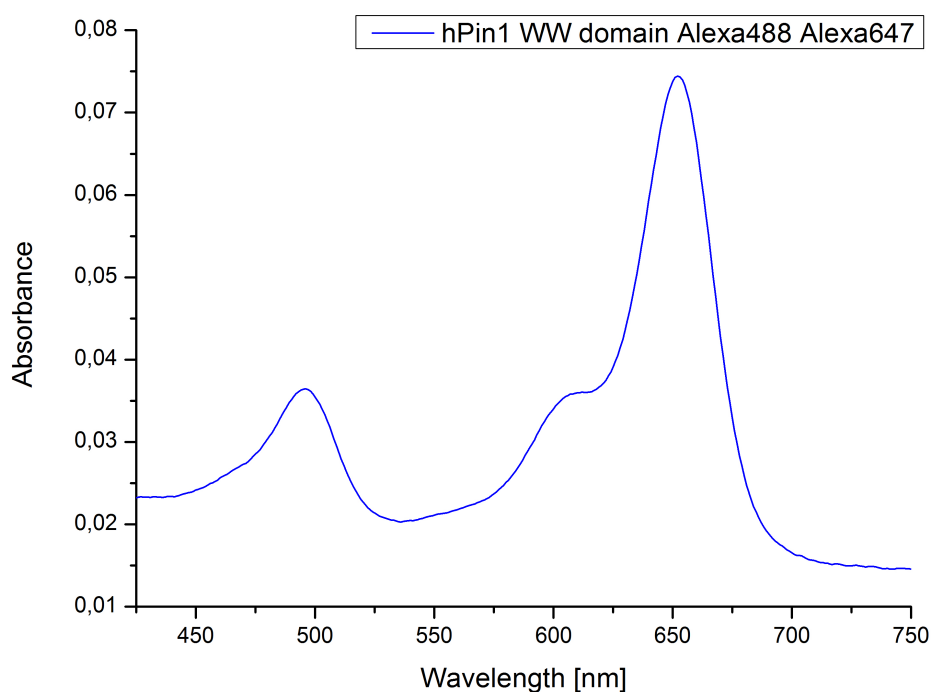


Figure 47: Spectra doubly labeled hPin1-WW-domain Alexa 488 Alexa 647

The concentration of Alexa Fluor 488 dye in the sample was 298 nM, while that of Alexa Fluor 647 was 221 nM. This yielded a labeling ratio of 1.34 : 1, which proved that there still was some free Alexa Fluor 488 dye left in the sample.

3.4 Discussion

The concept was to purify the expressed proteins by a His-tag. This procedure of protein purification is a well-established method (Hengen *et al.*, 1995^[29]). It had been used in our lab to purify several other proteins before. Besides, the WW-domain is known to be a stable protein. Consequently, the necessity of a fusion protein for the expression of the WW-domain had not been obvious. Screening of different expression conditions was time-consuming but still had to be done in order to optimize the expression of the WW-domains. The plasmid vectors used for the expression were changed so as to make sure that the lack of expression was not based on the used plasmid vector. So far, the expression of four different WW-domains had not worked. If the expression level was in fact strongly dependent on the kind of WW-domain used, one would expect at least some protein expression from one of the used WW-domains. Therefore, it seemed likely that the problem could be solved using a fusion-protein in order to support the WW-domain.

Several expression conditions like temperature and time but also bacterial strains and expression vectors had been changed up to this point. The next concept was to prepare a fusion-protein to prevent the WW-domain from being degraded in the bacterial cell or from precipitation during purification. Calmodulin was selected to be fused to the WW-domain because our lab already had experience with the purification of calmodulin and it was known to have a high expression level. Only one of the WW-domains was used for the fusion-protein as proof of principle. The results of the fusion-protein preparation showed that there was indeed expressed protein which had been purified by His-tag. However, purified protein in the SDS gel contained calmodulin only without a WW-domain fused to it.

The next strategy involved using both, an alternative affinity-tag, i.e. a GST fusion-protein, and another type of WW-domain, namely the hPin1-WW-domain. The hPin1-WW-expression vector was prepared according to Jäger *et al.* (Jäger *et al.*, 2001^[10]). It was crucial to use high amounts of protein for the purification and labeling of the GST-WW-domain fusion protein. This was necessary to compensate for the loss of sample in each of the nine purification steps.

The result was that the FRET sample contained free donor dye after the last purification step using a PD10 column. In each purification step, between 50 % to 80 % of the sample were lost. To prevent the loss of the sample, no further purification step was conducted.

4 Results: Measurement of WW-domain FRET sample

This chapter describes the results of the measurements of the WW-domain by several techniques like CD spectroscopy, DSC, and FRET in bulky solutions as well as single molecule FRET experiments.

4.1 Circular dichroism (CD)

Circular dichroism (CD) spectroscopy was used to analyze the secondary structure of the hPin1-WW-domain. The unlabeled hPin1-WW-domain and the singly Alexa 647 labeled hPin1-WW-domain were measured to compare their spectra to each other. Whereas the former was measured in a concentration of 35 μM in PBS, the concentration of the latter was 32 μM in PBS. The spectra of the unlabeled protein (green) and the singly-labeled one (blue) are shown in figure (48):

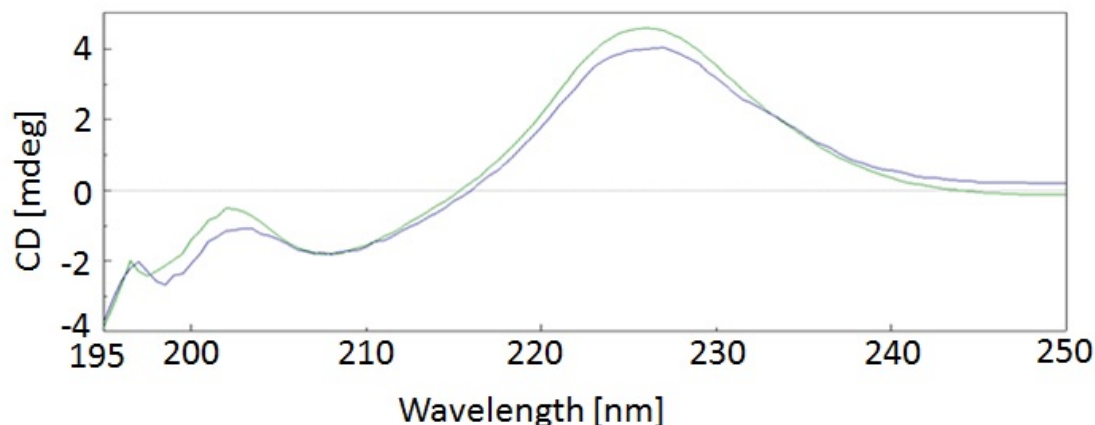


Figure 48: CD spectra of unlabeled (green) and singly-Alexa 647-labeled (blue) hPin1-WW-domain

Both spectra show a maximum molar ellipticity at 226 nm. This maximum molar ellipticity is characteristic for CD spectra of the hPin1-WW-domain, as shown by Jäger *et al.* (2009^[30]). However, the spectra of the hPin1-WW-domain is not typical for β -sheets. As a matter of fact, both spectra dealt with in this thesis are comparable to the CD spectra of the hPin1-WW-domain from that publication. A measurement of the FRET sample from the hPin1-WW-domain could not be recorded due to the fact that concentrations of the sample were too low. Given the fact that guanidine chloride absorbs light in the spectral range of the CD measurements, it was

not possible to record CD spectra of the unfolded hPin1-WW-domain.

4.2 Differential scanning calorimetry (DSC) of the hPin1-WW-domain.

The heat capacity C_P of the hPin1-WW-domain was measured using DSC. It was necessary to employ high concentrations of hPin1-WW-domain for the DSC measurements in order to get a detectable signal. Having been concentrated from a 20 μM sample of hPin1-WW-domain using a vacuum centrifugation (Eppendorf, Hamburg, Germany), the measured sample of unlabeled hPin1-WW-domain had a concentration of 660 μM in PBS. The reference buffer had been concentrated in the same way. The result of the measurement is shown in figure (49).

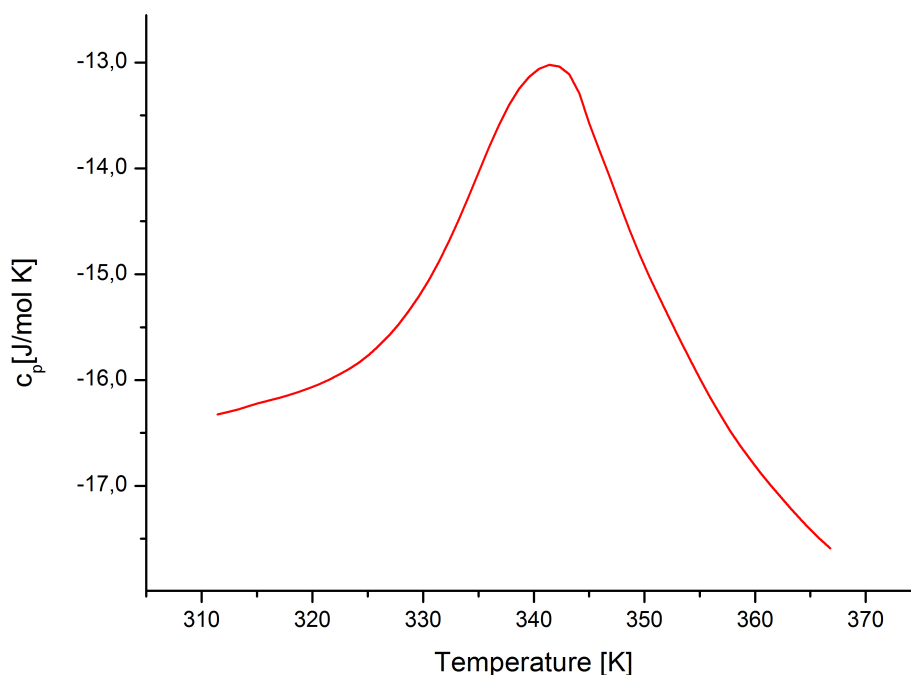


Figure 49: DSC measurement of the unlabeled hPin1-WW-domain.

The measurement showed a peak corresponding to the phase-transition of the hPin1-WW-domain at 342 K (69°C). The molar heat capacity C_P of the hPin1-WW-domain had a value of 3.3 J/mol · K (Ferguson *et al.*, 2001^[31]). Since the concentration of the labeled hPin1-WW-domain was too low, it was impossible to perform DSC measurements.

4.3 FRET experiments of hPin1-WW-domain in bulky solutions

The FRET sample hPin1-WW-domain was measured in a fluorescence photometer (FP). Besides, the dependence of protein unfolding on the guanidinium chloride concentration was determined by a concentration series. To do so, PBS was used for the measurement, with guanidinium chloride in concentrations between 0 M and 5.5 M. The hPin1-WW-domain FRET sample was added to the buffer in a concentration of 150 nM. Recording of the spectra was conducted in a range between 500 nm and 800 nm, while the sample was excited at 490 nm. The hPin1-WW-domain donor only sample was measured under the same buffer conditions. Figure (50) shows the spectra of the hPin1-WW-domain FRET sample (solid lines) measured under guanidinium chloride concentrations between 0 M and 5 M and the corresponding donor-only measurement (dashed lines):

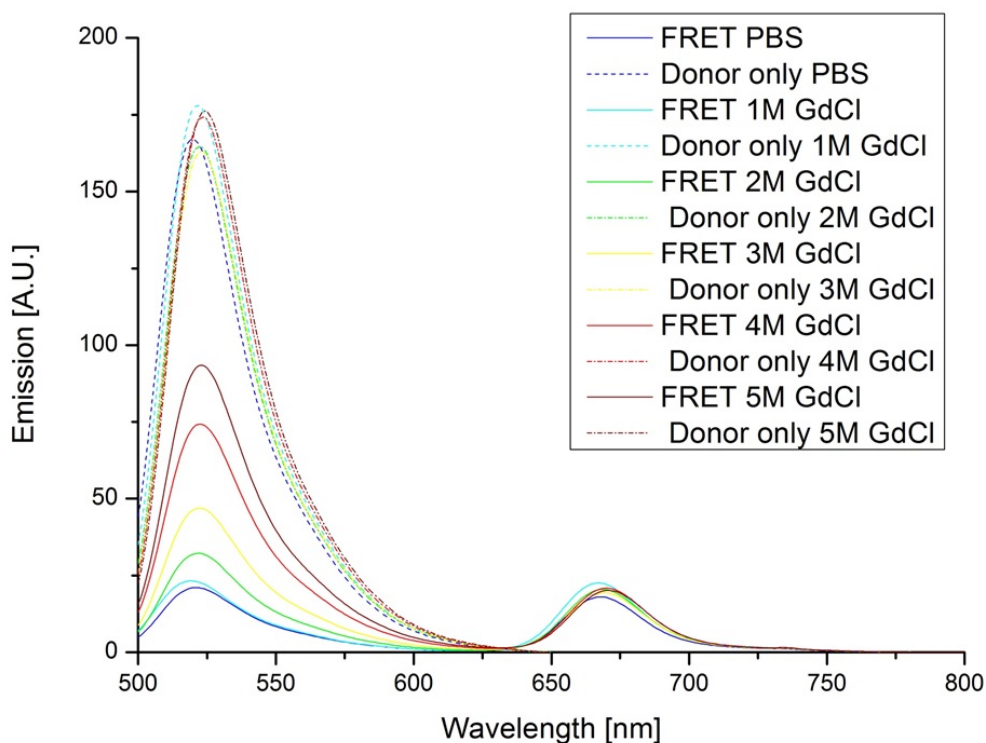


Figure 50: Emission spectra of hPin1-WW-domain measured in FP. Shown are spectra of the hPin1-WW-domain (solid lines) measured in PBS containing guanidinium chloride between 0 M and 5 M and the corresponding donor only samples (dashed lines) measured in the same buffer.

The spectra showed one intensity maximum at 520 nm corresponding to the emission of the donor dye (Alexa Fluor 488) and one intensity maximum at 667 nm which

accorded to the emission of the acceptor dye (Alexa Fluor 647). In order to calculate FRET efficiencies, it was necessary to measure a donor-only sample under the same conditions.

Intensity values for the calculation of the FRET efficiency E were determined using the maximum emission, according to equation (41):

$$E = 1 - \frac{F_{DA}}{F_D} \quad (41)$$

For the guanidinium chloride concentration series each buffer condition was measured 5 times. Guanidinium chloride was increased in 0.5 M steps. The dependence of unfolding on the hPin1-WW-domain is shown in the following figure (51):

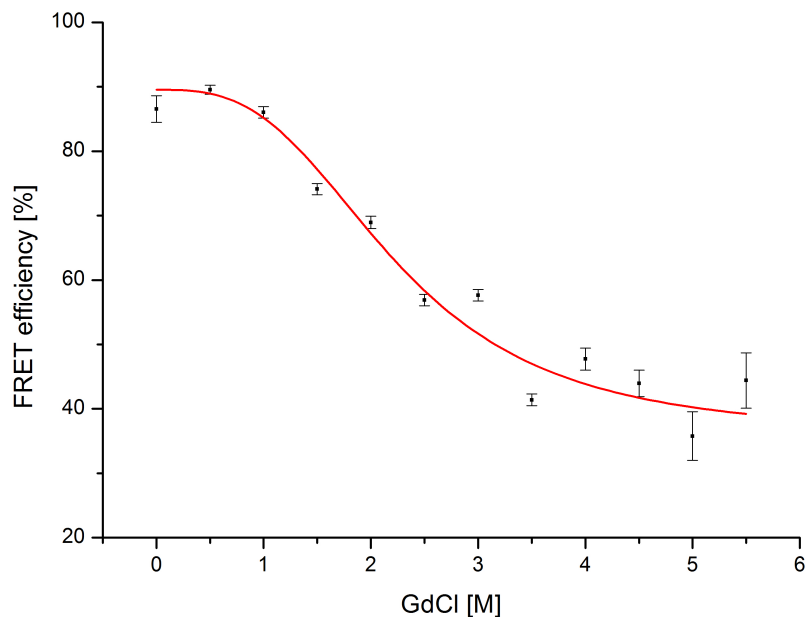


Figure 51: Evaluation of hPin1-WW-domain unfolding series. The hPin1-WW-domain was measured in PBS containing guanidinium chloride between 0 M and 5.5 M.

The figure shows lower FRET efficiencies for higher guanidinium chloride concentrations. Lower FRET efficiencies arise from the unfolding of the protein. Moreover, the data was fitted by a Hill fit (red). The Hill fit was achieved using equation (42).

$$y = \frac{y_{max} + (y_{min} - y_{max}) \cdot x^n}{(k^n + x^n)} \quad (42)$$

Here, y_{max} denotes the maximum FRET efficiency and y_{min} denotes the minimum FRET efficiency. The Hill fit delivered a n value of 3. For the binding constant k the

result was 2.24 M with an error of 0.08 M. This k value determines the guanidinium concentration where the chance for the hPin1-WW-domain being in the unfolded state is at 50 %.

4.4 Single molecule FRET experiments of the hPin1-WW-domain

The single molecule FRET measurements of the hPin1-WW-domain were conducted with different guanidinium chloride concentrations between 0 M and 6 M. Evaluation of these measurements has been given in chapter 2.13.8. Figure (52) shows the measurement evaluation of the hPin1-WW-domain in PBS.

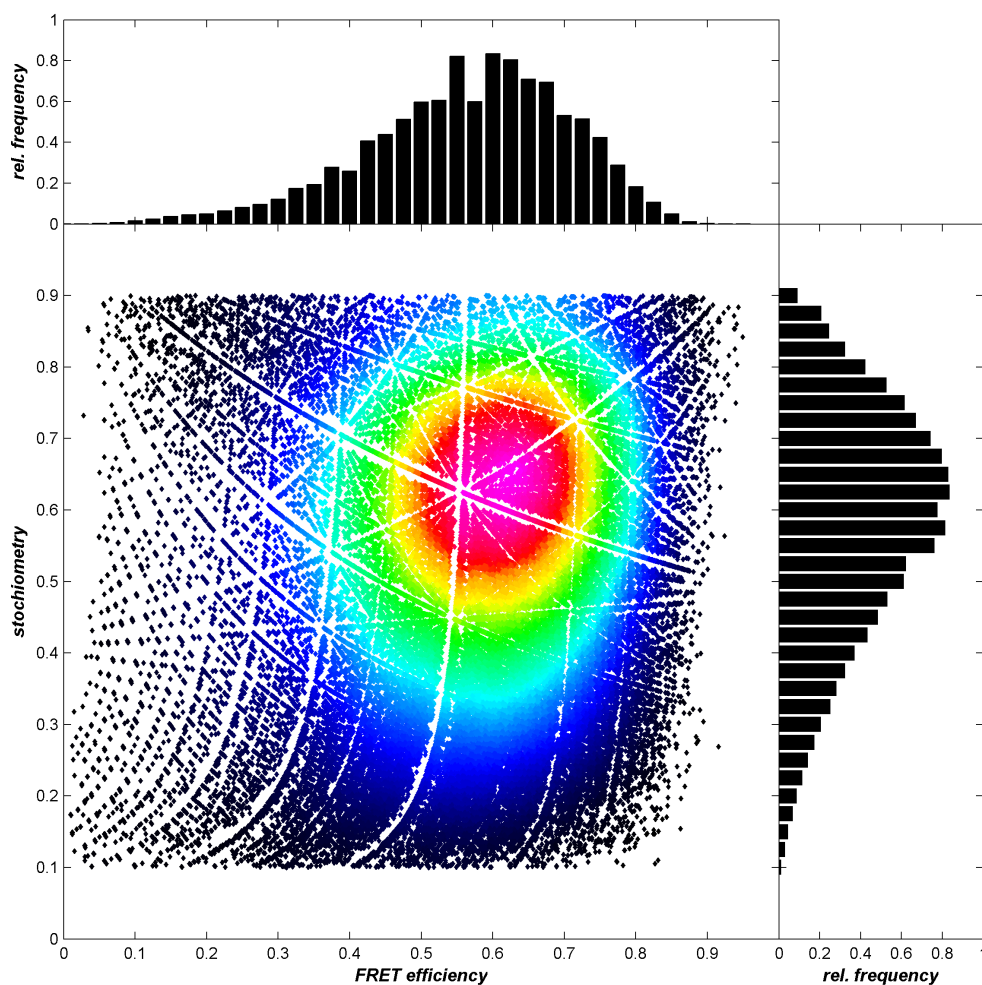


Figure 52: Evaluation of sm FRET measurement of hPin1-WW-domain in PBS. The 2D plot shows the FRET value and the stochiometry value of the measured sample.

In PBS, the hPin1-WW-domain is natively folded. FRET efficiency from the measurement was at around 60 %. The stoichiometry factor, which corresponds to the molar ratio of donor molecule and acceptor molecule to each other, was at around 0.6. This indicated that there was slightly more donor dye than acceptor dye present in the sample. A value of 0.5 would have indicated that donor dye and acceptor dye were present in equal amounts.

The hPin1-WW-domain measured in 6 M GdCl/PBS buffer had a FRET efficiency of around 25 % which showed that the protein was unfolded (shown in figure 53). The stoichiometry factor in this measurement was around 0.55.

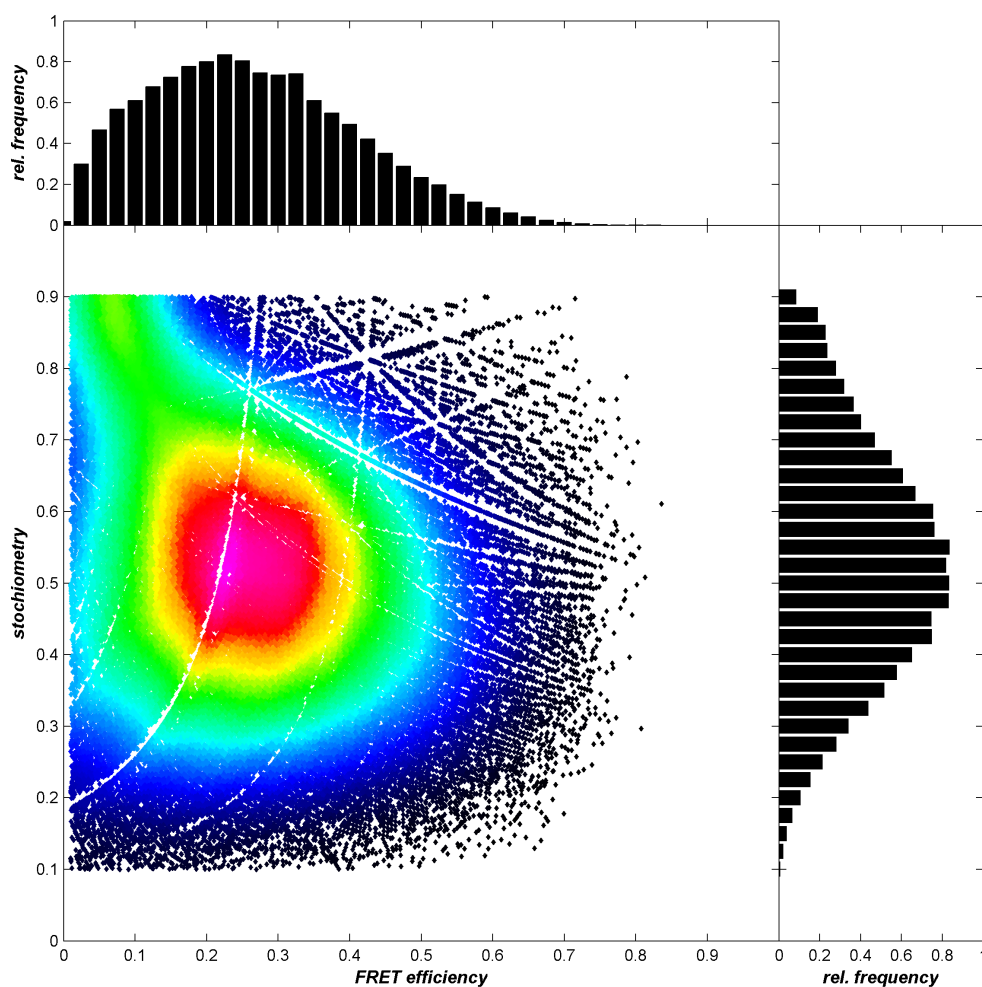


Figure 53: Evaluation of sm FRET measurement of hPin1-WW-domain in PBS containing 6 M guanidinium chloride. The 2D plot shows the FRET value and the stoichiometry value of the measured sample.

4.5 Nano-containers

Polymerosomes were filled with Alexa Fluor 647 in order to quantify their binding to the glass surface. The dye molecules inside the nano-containers were excited by laserlight at 635 nm with 1 μ W power. The scan of the glass surface was conducted using a confocal microscope (shown in figure (54)):

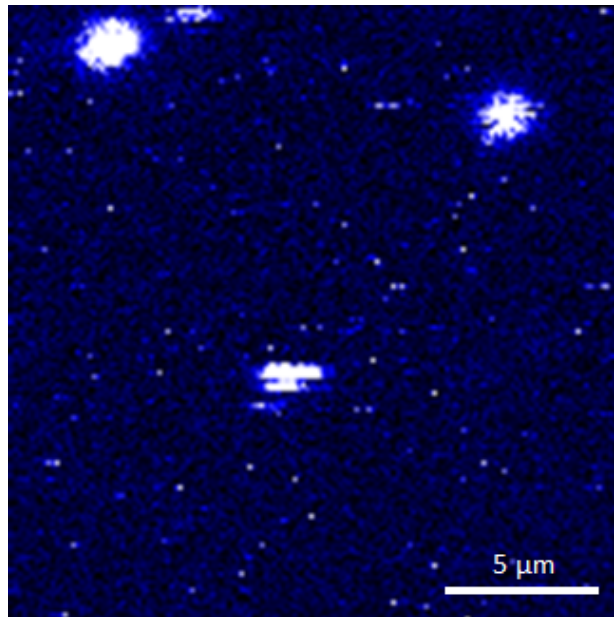


Figure 54: *Nano-containers with encapsulated Alexa Fluor 647 fixed on a glass surface. The scanned surface area had a size of 20 μ m x 20 μ m.*

In fact, the scanned surface showed three polymerosomes encapsulating the Alexa Fluor 647 dye which had a concentration of 1 μ M. Generally, the number of dye molecules encapsulated in one polymerosome can be estimated by the volume of a polymerosome and its concentration. In this case, the diameter of the polymerosomes was estimated from the extrusion-membrane size to be 100 nm. Consequently, the solvent volume inside the polymerosomes was estimated to a diameter of 50 nm which led to a calculated volume of $5.2 \cdot 10^{-19}$ liter. The amount of substance encapsulated in one polymerosome was calculated using equation (43):

$$n = c_{\text{dye}} \cdot V = 10^{-6} \text{ mol/l} \cdot 2.68 \cdot 10^{-19} \text{ l} = 2.68 \cdot 10^{-25} \text{ mol} \quad (43)$$

In equation (43), n is the amount of substance, c is the concentration of dye and V is the volume. The amount of substance was $2.68 \cdot 10^{-25}$ mol.

In a next step, the number of dye molecules present in one polymerosome was calculated using equation (44):

$$N = n \cdot N_A = 2.68 \cdot 10^{-25} \text{ mol} \cdot 6.022 \cdot 10^{25} \text{ mol}^{-1} = 0.44 \quad (44)$$

In equation (44), N_A is the Avogadro constant and N is the number of molecules. Calculation of the number of dye molecules in one polymerosome led to a value of 0.44, which suggested that, on average, there was one dye molecule encapsulated in every second polymerosome. This result is an estimation which has not been verified experimentally.

Nonetheless, these findings demonstrate the potential of this method for encapsulating molecules into nano-containers and thereby avoiding molecule diffusion out of focus during measurements. In the following, the hPin1-WW-domain FRET sample was encapsulated into polymerosomes attached to the surface. These results showed the potential of the encapsulation of molecules into polymerosomes.

4.6 Discussion

CD spectra of the hPin1-WW-domain were recorded so as to compare them to the published CD spectra of the protein. Thus, it was possible to show that the protein was folded in its native structure. Measured CD spectra from the labeled and unlabeled hPin1-WW-domain confirmed the CD spectra published by Jäger *et al.* (2009^[30]). The fact that the CD spectra of the labeled protein matched the published structure indicates that the native structure of the protein is not influenced by the attached fluorophores, a matter which is not self-evident.

The DSC measurement of hPin1-WW-domain was repeated five times. In all of these experiments, the protein did not fold back into its native structure after it had been heated up to 368 K (95°C). One might expect that small globular proteins are able to refold spontaneously. However, as it did not do so, the idea was that the degradation of the protein might have been based on precipitation. This precipitation could have occurred due to high protein concentrations in the sample which, however, were necessary in order to get detectable signal from the DSC measurement. Yet, the measured molar heat capacity C_P value of 3.3 J/mol · K did not agree with the value published in Ferguson *et al.* (2001^[31]), which was two times lower. In that publication, another type of WW-domain, i.e. the YAP65-WW-domain, was studied. Ultimately, this might have been the reason for the differences in measured molar heat capacity.

An important result of the FRET measurements was the dye-to-dye distances which were calculated. For the folded hPin1-WW-domain, the dye-to-dye distance was 4 nm, while in the unfolded state the dye-to-dye distance was 6 nm. In the crystal structure, the C-terminal-end and the N-terminal-end were 1 nm apart from each other. Both ends were flexible linker regions and the dye molecules were connected to them by C-5 linkers which had a length of roughly 1 nm. Therefore, the distance of 4 nm in the measurement fitted the expected distance. Notwithstanding, it is difficult to estimate the end-to-end distance for the unfolded protein from models. The Gaussian chain model is the commonly used model to calculate the end-to-end distance of completely unfolded proteins. Regarding protein unfolding using 6 M guanidinium chloride, it is known that salt bridges are removed while some H-bonds might be unaffected. Therefore, one cannot be sure whether the protein structure is entirely unfolded or still partially stabilized. Using the Gaussian chain

model, the dye- to- dye distance of the unfolded hPin1-WW-domain was calculated to be at around 7 nm. As had been expected, the measured dye-to-dye distance of 6 nm was below this calculated distance. Chung *et al.* (2012^[11]) published smFRET data of the FBP-WW-domain. In their work, the same FRET pair was used what makes their data comparable to our data. They measured a transfer efficiency of around 90 % for the folded WW-domain, which is comparable to our value of 88 %. For the unfolded state, they obtained a transfer efficiency of around 55 % -again comparable to our results of 50 %. Calculating dye-to-dye distances from the smFRET measurements lead to a resulting distance of 5.2 nm for the folded hPin1-WW-domain and 6.7 nm for the unfolded protein. The FRET efficiency values from the smFRET experiments cannot be compared completely to the bulk FRET measurements. It is most probable that this difference is an artifact arising from the data evaluation of the smFRET measurements. Other groups have solved this problem by applying smFRET measurements for molecules of a known distance r . Schuler *et al.* (2005^[32]) used poly-proline FRET samples of different lengths for a calibration of the setup. With regard to this calibration, it should be mentioned that labeled DNA molecules are another type of molecule that can be used for this.

As a result, the surface scan showed that the encapsulation of the dye molecules was successful. This proof of the method demonstrated the possibility to encapsulate molecules in the polymerosomes that we had produced. The correct concentration of protein for the encapsulation had to be determined in a next step. It was shown by Rigler *et al.* (2006^[33]), that the initial concentration of protein corresponds well to the concentration of encapsulated proteins. One method to get information from single molecules is to observe bleaching steps from the measured polymerosome, as shown by Rosenkranz *et al.* (2009^[34]).

5 Conclusion and outlook

The folding of proteins into a three-dimensional structure is in some cases a spontaneous process during which the protein adopts a state of the lowest free energy possible. The three-dimensional structure is stabilized by hydrophobic interactions as well as hydrogen bonds and salt bridges. The amount of energy needed to unfold a protein is dependent on the quality and the quantity of these stabilizations. Deducing fundamental findings about stability, structure and dynamics of proteins on the basis of their primary structure constitutes a milestone of science that might enable us to create new therapy methods against many diseases.

In this thesis, different methods were applied to examine stability and structure of the hPin1-WW-domain. It was possible to compare the values resulting from this analysis to those published in various scientific studies. In the course of the thesis work, the stability of the protein hPin1-WW-domain was determined using differential scanning calorimetry, whereas its structure was examined with the help of circular dichroism. However, the key experiments of this thesis are single molecule FRET measurements. smFRET experiments constitute a possibility to measure the folding and unfolding of the hPin1-WW-domain in detail and thus help to better understand the mechanisms of protein folding, especially for smaller proteins. In order to carry out those experiments, a FRET sample of the hPin1-WW-domain was produced.

A combination of several biochemical methods was necessary for the preparation of the FRET sample. The main difficulty of the FRET labeling was the necessity of nine purification steps which caused a high loss of sample. The preparation of the FRET samples required several changes in strategy for expression, purification and labeling.

In the end, we showed a way of expression and of purification applicable to small proteins like the WW-domain.

In summary, this thesis lays the ground for the preparation of FRET samples from small proteins. The thesis shows how these samples were analyzed in highly precise single molecule FRET measurements to determine their folding and unfolding.

Preparing the FRET sample was the most important step towards performing smFRET measurements. For future work, it would be interesting to determine the unfolding of the hPin1-WW-domain by smFRET measurements in order to define a midpoint of guanidinium concentration necessary for the unfolding of the protein. At this midpoint, the protein has the highest probability to change between the folded and unfolded state in a dynamic balance. Measuring the protein in a guanidinium concentration corresponding to the midpoint would give the opportunity to monitor the maximum number of unfolding and folding events. Thereby, it might be possible to resolve the trajectory of protein folding in more detail than has been the case so far. For these measurements, it could be helpful to use the polyerosome nanocontainers described in this thesis to keep the molecules in the laser focus. This will allow to measure the same molecule for a long time which is favourable in respect to signal to noise ratios.

For a detailed study of the folding mechanism, it might be advisable to have a slowly folding protein in order to be able to detect it. Hence, it might be necessary to prepare mutants of the WW-domain with a slower folding velocity. Liu *et al.* (2008^[35]) showed that mutagenesis of two to five amino acids in the hPin1-WW-domain can change its folding velocity by almost a factor of ten. This mutation might also be of use for further studies of WW-domain folding.

Appendix

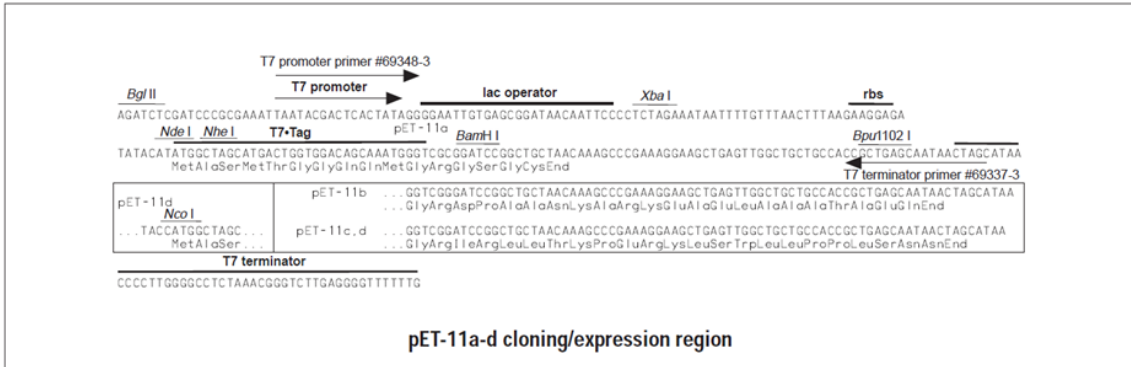


Figure 55: Vectormap pET11a (taken from Novagen, Madison, USA)

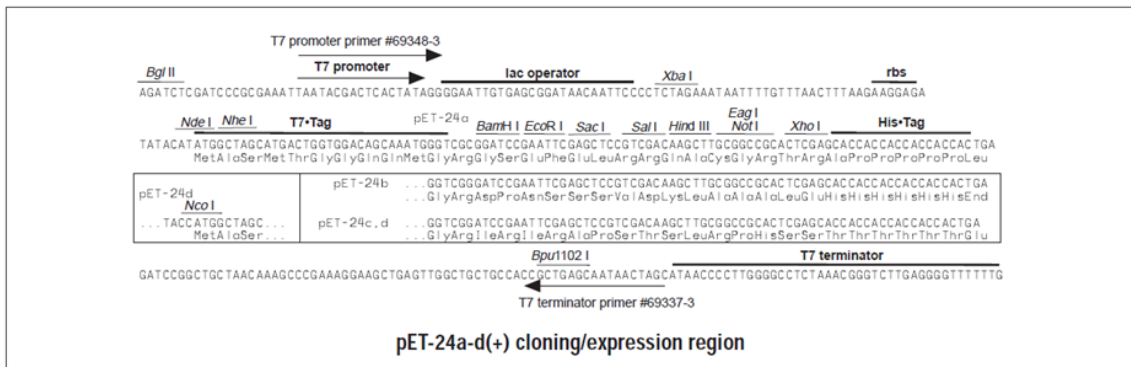


Figure 56: Vectormap pET24b (taken from Novagen, Madison, USA)

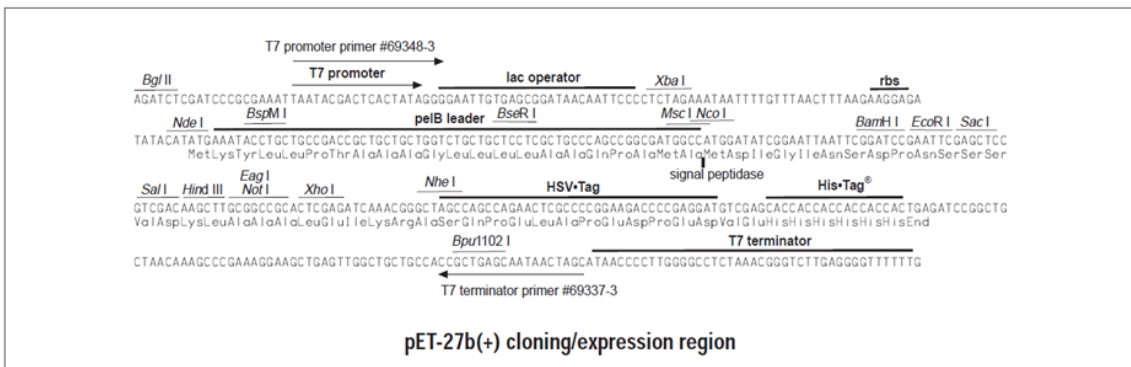


Figure 57: Vectormap pET27b (taken from Novagen, Madison, USA)

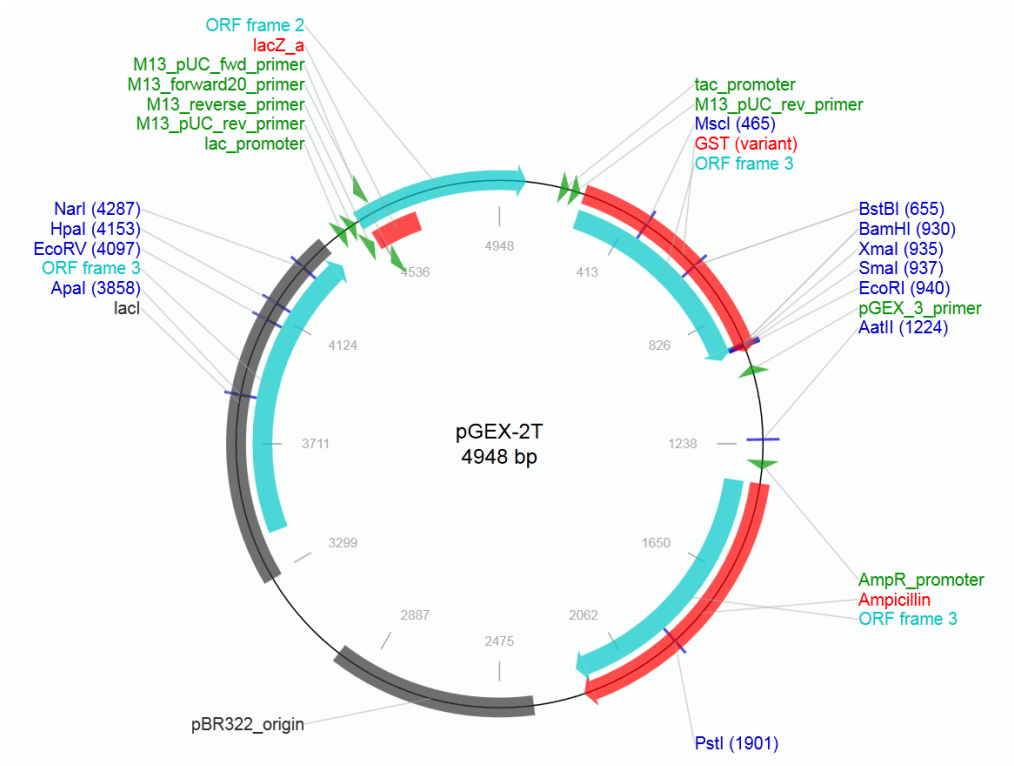


Figure 58: Vectormap of *pGEX2T* plasmid vector

Calmodulin-WW-domain fusion protein:

Amino acid sequence:

```

HHHHHHIEGRMADQLTEEQIAEFKEAFSLFDKD -
GDGTITTKELGTVMRSLGQNPTEAELQDMINEV -
DADGNGTIDFPEFLTMMARKMKDSTDSEEEIREA -
FRVFDKDGNGYISAAELRHVMTNLGEKLTDEEV -
DEMIREADIDGGQVNYEEFVQMMTAKGSDDDD -
DKCVPLPAGWEMAKTSSGQRYFLNHIDQTTTW -
QDPRKAMC

```

DNA sequence:

```

CATACCATCACCATCATATTGAAGGCCGATGGCTGATCAGCTGACTGAAGAGCAGATCGCA
GAGTTC AAGGAGGCCCTTCTCCCTCTTCGACAAAGACGGGGACGGCACCATCACCAAGGA
GCTGGGGACAGTGATGAGGTCTCTGGGACAGAACCCACTGAAGCCGAGCTGCAGGACATG
ATCAACGAGGTGGACGACAGATGGGAACGGGACCATTGACTTCCCGAGTTCCTGACCATGAT
GGCCAGAAAGATGAAGGACACGGACAGCGAGGAGGAGATCCGAGAGGCCTTCGGTGTCTTT
GACAAGGACGGCAACGGGTACATCAGCGCCGAGAGCTGCGCCAGTCATGACGAACCTGG
GCGAGAAGCTGACCGACGAGGAGGTGGACGAGATGATCAGGGAGGCCGACATTGACGGGG
ACGGCCAGGTCAATTATGAAGAGTTTGTACAATGATGACTGCAAAAGGATCCGACGATGACGA
TAAATGTGTTCCGCTGCCAGCAGGTTGGGAAATGGCAAAAACGACGACGGTTCAGCGTTATT
TTCTGAATCATATTGATCAGACCACCCTGGCAGGACCCTCGTAAAGCAATGTGT

```

His-tag
 Calmodulin
 Factor xa rec. Site
 WW-domain

Figure 59: Amino acid and DNA sequence of Calmodulin-WW-domain

Bibliography

- [1] ishbytes.blogspot.com, <http://ishbytes.blogspot.com> (2013).
- [2] C. Branden, J. Tooze, *Introduction to protein structure*, Garland Science New York (1999).
- [3] J. Buchner, T. Kieferhaber, *Protein Folding Handbook, Volume 1*, Wiley-VCH Weinheim (2005).
- [4] J. C. Kendrew, “A three-dimensional model of the myoglobin molecule obtained by x-ray analysis” *Nature* (1958) **181**(4610), 662.
- [5] C. B. Anfinsen, “Principles that Govern the Folding of Protein Chains” *Science* (1973) **181**(4096), 223–230.
- [6] <http://jonlieffmd.com>, <http://jonlieffmd.com> (2013).
- [7] A. Fersht, *Structure and mechanism in protein science*, Freeman New York (2002).
- [8] O. Staub, D. Rotin, “WW domains” *Structure* (1996) **4**, 495–499.
- [9] PDB data bank, “PDB statistics”, http://pdbeta.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html (2013).
- [10] M. Jäger, H. Nguyen, J. Crane, J. Kelly, M. Gruebele, “The folding mechanism of a beta-sheet: the WW domain” *Journal of molecular biology* (2001) **311**, 373–393.
- [11] H. Chung, K. McHale, J. Louis, W. Eaton, “Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times” *Science* (2012) **335**, 981–984.
- [12] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, Erlich H., “Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.” *Cold Spring Harb. Symp. Quant. Biol* (1986) **51**, 263–273.
- [13] S. C. Gill, P. H. von Hippel, “Calculation of protein extinction coefficients from amino acid sequence data.” *Analytical Biochemistry* (1989) **182**, 319–326.

- [14] ExPASy, “ExPASy protein parameter”, <http://web.expasy.org/cgi-bin/protparam/protparam> (2013).
- [15] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, “Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels” *Analytical Chemistry* (1996) **68**, 850–858.
- [16] Norma J. Greenfield, “Using circular dichroism spectra to estimate protein secondary structure” *Nat. Protocols* (2007) **1**(6), 2876–2890, 10.1038/nprot.2006.202.
- [17] N. Sreerama, R. Woody, “Computation and Analysis of Protein Circular Dichroism Spectra” *Methods in enzymology* (2004) **383**, 318–351.
- [18] M. J. O’Neill, “The Analysis of a Temperature-Controlled Scanning Calorimeter” *Analytical Chemistry* (1964) **36**(7), 1238–1245.
- [19] T Förster, “Zwischenmolekulare Energiewanderung und Fluoreszenz” *Annalen Der Physik* (1948) **437**(1-2), 55–75.
- [20] Haugland Stryer, “Energy transfer: a spectroscopic ruler.” *PNAS* (1967) **58**(2), 719–726.
- [21] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, S. Weiss, “Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor” *PNAS* (1996) **93**, 6264–6268, 1658.
- [22] B.K. Mueller, E. Zaychikov, C. Braeuchle, D.C. Lamb, “Pulsed Interleaved Excitation” *Biophysical Journal* (2005) **89**, 3508–3522.
- [23] C. Eggeling, S. Berger, L. Brand, J. R. Fries, J. Schaffer, A. Volkmer, C. A. M. Seidel, “Data registration and selective single-molecule analysis using multiparameter fluorescence detection” *J Biotechnol* (2001) **86**, 163–180, 2372.
- [24] Volodymyr Kudryavtsev, Martin Sikor, Stanislav Kalinin, Dejana Mokranjac, Claus A. M. Seidel, Don C. Lamb, “Combining MFD and PIE for Accurate Single-Pair Förster Resonance Energy Transfer Measurements” *Chemphyschem* (2012) **13**(4), 1060–1078.

- [25] A. A. Deniz, M. Dahan, J. R. Grunwell, T. Ha, A. E. Faulhaber, D. S. Chemla, S. Weiss, P. G. Schultz, “Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations” *PNAS* (1999) **96**, 3670–3675, 2117.
- [26] M. Macias, G. Virginie, C. oschkinat, H. Oschkinat, “Structural analysis of WW domains and design of a WW prototype” *Nature structural biology* (2000) **7**, 375–379.
- [27] F. Liu, “An experimental survey of the transition between two-state and downhill protein folding scenarios” *PNAS* (2007) **105**, 2369–2374.
- [28] J. Lakowicz, *Principles of Fluorescence Spectroscopy, Third Edition*, Springer Science+Business Media New York (2006).
- [29] P. Hengen, “Purification of His-Tag fusion proteins from *Escherichia coli*” *Trends in biochemical sciences* (1995) **20**, 285–286.
- [30] M. Jäger, M. Dendle, J. Kelly, “Sequence determinants of thermodynamic stability in a WW-domain an all beta-sheet protein” *Protein Science* (2009) **18**, 1806–1813.
- [31] N. Ferguson, C. Johnson, M. Macias, H. Oschkinat, A. Fersht, “Ultrafast folding of WW domains without structured aromatic clusters in the denatured state” *PNAS* (2001) **98**, 13002–13007.
- [32] B. Schuler, E. Lipman, P. Steinbach, M. Kumke, W. Eaton, “Polyproline and the spectroscopic ruler revisited with single-molecule fluorescence” *PNAS* (2005) **102**, 2754–2759.
- [33] P. Rigler, W. Meier, “Encapsulation of Fluorescent Molecules by Functionalized Polymeric Nanocontainers: Investigation by Confocal Fluorescence Imaging and Fluorescence Correlation Spectroscopy” *JACS* (2006) **128**, 367–373.
- [34] T. Rosenkranz, A. Katranidis, D. Atta, I. Gregor, M. Enderlein, J. Grzelakowski, P. Rigler, W. Meier, J. Fitter, “Observing Proteins as Single Molecules Encapsulated in Surface-Tethered Polymeric Nanocontainers” *CHEMBIOCHEM* (2009) **10**, 702–709.
- [35] F. Liu, D. Du, A. Fuller, J. Davore, P. Wpif, J. Kelly, M. Gruebele, “An experimental survey of the transition between two-state and downhill protein folding scenarios” *PNAS* (2008) **105**, 2369–2374.

List of Figures

1	Schematic drawing of an E. coli cell (taken from: ishbytes ([1])) .	3
2	Basic structure of amino acids (adapted from: Branden and Tooze ^[2])	4
3	Peptide bond of amino acids. Blue arrows indicate bond angles (adapted from: Branden and Tooze ([2]))	4
4	Primary and secondary structure of proteins (adapted from: Buchner and Kieferhaber ([3]))	5
5	Tertiary and quaternary structure of proteins (adapted from: Buchner and Kieferhaber ([3]))	6
6	Schematic illustration of the protein folding funnel (taken from: ([6]))	7
7	Schematic illustration of β-sheets in proteins (adapted from: Buchner and Kieferhaber ([3]))	9
8	Primary structure of the hPin1-WW-domain. Amino acids emphasized in red are inside the β -strands.	9
9	Hydrophobic clusters of hPin1-WW-domain. Amino acids in red and green form the hydrophobic clusters (adapted from: Protein data bank identification: Pin1 ([9]))	10
10	H-bond network of hPin1-WW-domain. H-bonds are shown in red while salt bridges are shown in yellow (adapted from: Protein data bank identification: Pin1 ([9]))	11
11	Scheme of nano-containers attached to the surface.	26
12	Schematic drawing of ABA tribloc co-polymer	27
13	Chamber preparation for polymerosome measurement.	28
14	D and L isomers of aminoacids	29
15	Circular dichroism. E_L is shown in blue, E_R is shown in black, the elliptical polarization is shown in red and angle theta is shown in green.	30
16	CD measurement evaluation. Shown is the measurement of the hPin1-WW-domain in PBS.	31
17	Schematic illustration of the DSC setup. Measurement chambers are shown in blue	33
18	Jablonski diagram	34
19	Stokes shift of Alexa Fluor 488. Absorption spectra of Alexa 488 in solid blue line, Emission spectra in dashed blue line.	36

20	Orientation factor κ^2. The transition dipole of donor (D) and acceptor (A) are shown in red.	39
21	Transfer efficiency dependent on r/R_0	40
22	Absorption spectra of Alexa 647 (red) and Alexa 750 (blue)	42
23	Normalized emission spectra of Alexa 647 (red) and Alexa 750 (blue)	42
24	Schematic illustration of maleimide labeling reaction. The sulfur is emphasized in yellow.	43
25	Pulsed interleaved excitation (PIE).	46
26	Direct acceptor excitation Shown are the absorption spectra of Alexa 488 (blue) and Alexa 647 (red) and the lasers used for excitation.	47
27	Spectral crosstalk. Shown are the emission spectra of Alexa 488 (blue) and Alexa 647 (red) and the properties of the emission filters (dashed lines).	48
28	Microtime200 setup.	50
29	Sequence alignment of WW-domains. Shown are the amino acid sequences of the Yes kinase associated protein 65 WW-domain from human (YAP65-WW-domain), the formin binding protein 28 from mouse (FBP28-WW-domain), the YJQ8-WW-domain from yeast and the prototype-WW-domain which is an artificial protein.	53
30	Restriction scheme of WW-domain genes for separation.	54
31	WW-domain genes after restriction in agarose gel stained by ethidium bromide.	55
32	Scheme of WW-domain proteins.	56
33	Schematic illustration of the calmodulin-YAP65-WW-domain.	60
34	Ligation scheme of the calmodulin-WW-domain into the expression vector.	61
35	Expression vector of the calmodulin-WW-domain fusion protein.	61
36	SDS gel of the Calmodulin-WW-domain fusion protein. The first line shows the protein ladder, the second line shows the flow through fraction and the third line shows the eluat which contains the calmodulin.	62

37	Sequence alignment of hPin1-WW-domain and other WW-domains. Shown are the amino acid sequences of the Yes kinase associated protein 65 WW-domain from human (YAP65-WW-domain), the formin binding protein 28 from mouse (FBP28-WW-domain), the YJQ8-WW-domain from yeast, the prototype-WW-domain which is an artificial protein and the rotamase WW-domain from human (hPin1-WW-domain)	63
38	hPin1-WW-domain in pGEX2T vector with QC primers.	65
39	hPin1-WW-domain in pGEX2T with N- and C-terminal cysteins.	65
40	SDS-PAGE of GST-hPin1-WW-domain cut by thrombin. The first line shows the protein ladder, the lines two to five contains the GST-hPin1-WW-domain cut by thrombin with incubation times between 0 h and 9 h.	66
41	hPin1-WW-domain purification	67
42	Identification of Fraction 2	68
43	Identification of Fraction 3	69
44	Determination of concentration of hPin1-WW-domain.	70
45	Ion exchange chromatography of hPin1-WW-domain Alexa Fluor 647	72
46	hPin1-WW-domain Alexa Fluor 647 spectrum	73
47	Spectra doubly labeled hPin1-WW-domain Alexa 488 Alexa 647	75
48	CD spectra of unlabeled (green) and singly-Alexa 647-labeled (blue) hPin1-WW-domain	77
49	DSC measurement of the unlabeled hPin1-WW-domain.	78
50	Emission spectra of hPin1-WW-domain measured in FP. Shown are spectra of the hPin1-WW-domain (solid lines) measured in PBS containing guanidinium chloride between 0 M and 5 M and the corresponding donor only samples (dashed lines) measured in the same buffer.	79
51	Evaluation of hPin1-WW-domain unfolding series. The hPin1-WW-domain was measured in PBS containing guanidinium chloride between 0 M and 5.5 M.	80

52	Evaluation of sm FRET measurement of hPin1-WW-domain in PBS. The 2D plot shows the FRET value and the stoichiometry value of the measured sample.	81
53	Evaluation of sm FRET measurement of hPin1-WW-domain in PBS containing 6 M guanidinium chloride. The 2D plot shows the FRET value and the stoichiometry value of the measured sample.	82
54	Nano-containers with encapsulated Alexa Fluor 647 fixed on a glass surface. The scanned surface area had a size of 20 μm x 20 μm	83
55	Vectormap pET11a (taken from Novagen, Madison, USA)	89
56	Vectormap pET24b (taken from Novagen, Madison, USA)	89
57	Vectormap pET27b (taken from Novagen, Madison, USA)	89
58	Vectormap of pGEX2T plasmid vector	90
59	Amino acid and DNA sequence of Calmodulin-WW-domain	90

Acronyms

ADB	Agarose dissolving buffer
APS	Ammonium persulfate
ATP	Adenosine triphosphate
bp	Basepairs
BSA	bovine serum albumin
CD	Circular dichroism
DETA	Diethylen triamin
Da	Dalton
DNA	Desoxyribonukleinsäure
dNTP	Desoxyribonukleotid
DSC	Differential scanning calorimetry
DTT	Dithiothreitol
EDTA	Ethylendiamintetraessigsäure
FP	Fluorescence photometer
FRET	Förster resonance energy transfer
Fwd	Forward
GdCl	Guanidinium chloride
GST	Glutathione S-transferase
HCl	Hydrochloric acid
HPLC	High-performance liquid chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
KOH	Potassiumium hydroxid

LB	Lysogenic broth
MALDI-TOF	Matrix-assisted laser desorption/ionization time of flight
MFD	Multi-parameter fluorescence detection
NA	Numerical aperture
NaCl	Sodium chloride
NaOH	Sodium hydroxide
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDMS	Polydimethylsiloxan
PEG	Polyethylene glycol
PIE	Pulsed interleaved excitation
QC	Quick change
Rev	Reverse
RNA	Ribonucleic acid
Rpm	Rounds per minute
RT	Room temperature
SDS	Sodium dodecyl sulfate
Sm	Single molecule
SPAD	single photon avalanche diode
TCSPC	time-correlated single photon counting
TRIS	Tris(hydroxymethyl)aminomethane
UV	Ultra violet

The common acronyms as suggested by the *Système International* were used in this work.

Nucleotides in nucleic acid sequences are named by their first letter (A=adenine; C=cytosine; G=guanine; T=thymine).

Amino acids are either abbreviated by their three or one letter acronyms:

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Acknowledgments

Ich möchte mich bei **Jörg Enderlein** für die Möglichkeit bedanken, meine Doktorarbeit in seiner Gruppe anzufertigen. Ich konnte sehr von deinen Erklärungen profitieren und viel über physikalische Zusammenhänge lernen.

Bei dem Mitgliedern meines Thesis Committees **Helmut Grubmüller** und **Andreas Janshoff** möchte ich mich für die hilfreichen und unterstützenden Diskussionen während meiner Doktorarbeit bedanken.

Von **Markus Jäger** habe ich viel über Protein-Aufreinigung und Protein-labeling gelernt, für all die hilfreichen Hinweise und Erklärungen möchte ich mich bedanken.

Ingo Gregor möchte ich ganz herzlich für seine Unterstützung am Microtime Setup und besonders für seine Hilfe bei physikalischen Fragen jeglicher Art danken!

Stefanie Kramer danke ich für ihre Unterstützung und die gute Atmosphäre im Labor. Unsere gemeinsame Zeit im Büro werde ich immer in guter Erinnerung behalten, vor allem, weil wir uns zwischendurch immer über die selben Dinge amüsieren konnten.

Qui Van danke ich für die Unterstützung bei biochemischen Fragestellungen im Labor.

Mira Prior war auf Tagungen und Ausflügen unsere "Enten-Mama", auf die man sich immer verlassen konnte. Sehr dankbar bin ich besonders für deine Hilfe beim Verstehen von physikalischen Themen und bei der Benutzung von LaTeX.

Kerstin Weiß musste während unserer gemeinsamen Promotionszeit wahrscheinlich am meisten unter meinen Scherzen "leiden". Ich habe aber natürlich nur versucht, dich auf lustige Weise reinzulegen um deinen Gefahreninstinkt zu verbessern! Wir haben selbstverständlich auch sinnvolle Dinge getan, wie z.B. das Aufspüren von Horcruxen. Und ich habe dir natürlich sehr viel zu verdanken, weil du mit deinem "FRET-Kuchen" meine ganze Arbeit erst möglich gemacht hast! Vielen Dank dafür nocheinmal!

"Wenn wir jetzt in einer Sitcom wären..." So haben viele alberne Ideen mit **Anja Huss** angefangen. Häufig wurden dabei unerklärliche Dinge dann aber auf das "Quantum entanglement" zurückgeführt!

Ich möchte allen Mitgliedern der AG Enderlein für die nette Zusammenarbeit danken! Dir, **Lara**, möchte ich dafür danken, dass du immer für mich da bist! Wenn mal etwas nicht geklappt hat, hast du mich aufgemuntert und wenn etwas gut lief, hast du dich mit mir zusammen gefreut! Wenn man jemanden wie Dich an seiner Seite hat, kann im Leben nichts schief gehen!

Meinen Eltern, **Brigitte und Gunther**, möchte ich ganz herzlich dafür danken, dass sie mich immer unterstützt haben und mir mein Studium ermöglicht haben. Schon als kleiner Junge wollte ich gerne "Forscher" werden. Ihr habt einen großen Teil dazu beigetragen, dass ich mir diesen Wunsch erfüllen konnte. Denn niemand weiß, ob ich ohne die Besuche im Museum bei den Dinosauriern oder auch die Diskussionen über "Knochen geformt wie ein Eimer" wirklich Naturwissenschaftler geworden wäre.

Meiner Schwester **Nina** möchte ich dafür danken, dass sie immer für mich da ist wenn ich sie brauche! Ich weiß nicht, ob es solche zwei Menschen nocheinmal gibt, die sich über Stunden mit so vielen Insidersprüchen unterhalten könnten, wie uns?! Ich finde, dass dies unsere tiefe Verbundenheit zueinander widerspiegelt.

Solange ich zurückdenken kann, habe ich meine **Oma** immer als einen Menschen erlebt, der für andere da ist. Besonders schön für mich war das natürlich als Kind, da ich nur ein paar Treppen zu dir laufen musste. Mir hat dieses Gefühl immer sehr viel Sicherheit gegeben!

Curriculum vitae

Phillip Gunther Kroehn

Date of birth: 03.März 1982

Place of birth: Hannover, Niedersachsen

Nationality: German

Education

- 10/2008 - 07/2009 Diploma in Biology, Georg-August University Goettingen:
Diploma thesis in Biophysics, Georg-August University Goettingen:
"Biophysikalische Charakterisierung der Motilität und des Umsatzes
chimärer Motorproteine - Fortgeschrittene
Fluoreszenzmikroskopie in der Biophysik"
Supervisor: Prof. Dr. Christoph F. Schmidt
- 10/2003 - 07/2009 Biology studies at Georg-August University Goettingen
- 08/1999 - 06/2002 High School - CJD Christophorusschule Elze

Publications

Conference contributions

Talks "Protein folding dynamics studied by single molecule FRET ", Symposium
SFB 755 - Nanoscale photonic imaging, 2013.

Posters "SMS - FRET spectroscopy has emerged as a versatile tool in life sciences",
16th International Workshop on Single Molecule Spectroscopy, Berlin, 2010.

"SMS - FRET spectroscopy has emerged as a versatile tool in life sciences", 17th
International Workshop on Single Molecule Spectroscopy, Berlin, 2011.

"SMS - FRET spectroscopy has emerged as a versatile tool in life sciences", 8th
International PhD student symposium Horizons in Molecular Biology, Goettingen,
2011.

"Protein folding dynamics studied by single molecule FRET", 18th International

Workshop on Single Molecule Spectroscopy, Berlin, 2012.

Göttingen, 13.08.2013